Open Access **Baghdad Science Journal** **P-ISSN: 2078-8665**
2020, 17(1) Supplement (March):361-370 **E-ISSN: 2411-7986**

DOI: https://dx.doi.org/10.21123/bsj.2020.17.1(Suppl.).0361

New Versions of Liu-type Estimator in Weighted and non-weighted Mixed Regression Model

Mustafa Ismaeel Naif Alheety

Received 26/9/2018, Accepted 19/9/2019, Published 18/3/2020

Abstract:

This paper considers and proposes new estimators that depend on the sample and on prior information in the case that they either are equally or are not equally important in the model. The prior information is described as linear stochastic restrictions. We study the properties and the performances of these estimators compared to other common estimators using the mean squared error as a criterion for the goodness of fit. A numerical example and a simulation study are proposed to explain the performance of the estimators.

Key words: Liu estimator, Mean squared error matrix, Mixed estimator, Stochastic restricted Liu estimator, Weighted mixed Liu estimator.
Mathematics Subject Classification Primary: 62J05; Secondary 62J07.

Introduction:

The linear regression model is given as the follows:

$$\underline{Y} = X\underline{\beta} + \underline{\epsilon}, \quad \dots(1)$$

where $\underline{Y}$ is an $n \times 1$ column of observations that explain the dependent variable, X is an $n \times p$ matrix of observations on p independent variables, $\underline{\beta}$ is a $p \times 1$ column of unknown parameters and $\underline{\epsilon}$ is an $n \times 1$ column of residuals, with an expected value equal to zero and a variance – covariance matrix equal to $\sigma^2 I_n$.
When all the assumptions of the linear model in (1) have been satisfied, the ordinary least squares estimator, denoted (OLS), will be the best linear unbiased estimator for (1) and is given as follows:

$$\underline{\hat{\beta}} = S^{-1}X'\underline{Y}, \quad \dots \quad (2)$$

where S = X′X. The OLS estimator is not always be a good estimator when the multicollinearity is present; consequently, the goodness of the OLS estimator will be missed. Neter, (1) said that in the process of fitting regression model, when one independent variable is nearly combination of other independent variables and this will affect parameter estimates. Multicollinearity may cause serious difficulties.

Department of Mathematics, College of Education for Pure Sciences, University of Anbar, Anbar, Iraq.
E-mail:eps.mustafa.ismaeel@uoanbar.edu.iq
*ORCID ID:0000-0002-1684-7682

Variances of parameter estimates may be unreasonably large, parameter estimates may not be significant and a parameter estimate may have a sign different from what is expected.

Thus, the detection of multicollinearity has to be made to reduce the effect on the estimation. The measures most applied to detect multicollinearity are the Variance Inflator Factor (VIF) and the Condition Number (CN) and the researchers are still working for this subject (2).

To reduce the effect of this problem, the biased estimation technique has been developed. Therefore, many new biased estimators have been proposed, such as the RR estimator (3). From the theory that the combination of two different estimators might inherit the advantages of both estimators, Liu (4) combined the Stein estimator with the RR estimator and proposed the Liu estimator (LE) as follows:

$$\underline{\hat{\beta}}_{LE}(d) = (S+I)^{-1}\left(X'\underline{Y} + d\underline{\hat{\beta}}\right), \quad \dots \quad (3)$$

where 0 < d < 1. Fela O. and Selahattin k. (5) discussed the predictive performance of the Liu estimator comparing it with ordinary least squares, principal components and Ridge regression estimators. Also Sivarajah and Pushpakanthie (6) proposeed a new biased estimator namely modified almost unbiased Liu estimator by combining almost unbiased Liu estimator (AULE) and ridge estimator (RE) in a linear regression model when multicollinearity is present among the independent variables.

In addition to the sample information, some exact or stochastic restrictions may be available for the unknown parameter of the model under consideration, then this will help to overcome the multicollinearity problem. Therefore, suppose that we have some prior information about β in the form of stochastic linear restrictions as follows:

$$\underline{h} = H\underline{\beta} + \underline{e} \sim (0, \sigma^2 V), \quad \ldots \quad (4)$$

where $\underline{h}$ is a $j \times 1$ matrix that may be interpreted as a random vector with $\mathrm{E}(\underline{h}) = \mathrm{H}\underline{\beta}$, H is a $j \times p$ known matrix and V is assumed to be a known and positive definite (pd). Additionally, it is assumed that $\underline{\epsilon}$ is stochastically independent of $\underline{e}$.

Theil and Goldberger (7,8) introduced the mixed estimation technique by unifying the sample and the prior information in equation (4) in a common model. These authors introduced what they called the ordinary mixed estimator (OME) as follows:

$$\underline{\hat{\beta}}_m = (S + H'V^{-1}H)^{-1}(X'\underline{Y} + H'V^{-1}\underline{h}) \ldots \quad (5)$$

To improve the performance of OME estimator, Hu Yang and Jianwen Xu (9) introduced what they called the stochastic restricted Liu estimator by combining the OME and LE as follows:

$$\underline{\hat{\beta}}_{SRLE}(d) = \underline{\hat{\beta}}(d) + S^{-1}H'(V + H\, S^{-1}H')^{-1}\left(\underline{h} - H\underline{\hat{\beta}}(d)\right) \quad \ldots \; (6)$$

Nilgün Yıldız (10) provided a new alternative estimator, the stochastic restricted Liu-type estimator, which is obtained by combining the OME and Liu-type estimator in the following way:

$$\underline{\hat{\beta}}_{\mathbf{SRLTE}}(k,d) = S_k^{-1}(S - dI)(S + H'V^{-1}H)^{-1}(X'\underline{Y} + H'V^{-1}\underline{h}) \; \ldots (7)$$

When the sample information given by (1) and the prior information presented by (4) are assigned as not equally important, Schffrin and Toutenburg (11) introduced the weighted mixed estimator (WME) as follows:

$$\underline{\hat{\beta}}_w = (S + wH'V^{-1}H)^{-1}(X'\underline{Y} + wH'V^{-1}\underline{h}), \quad \ldots(8)$$

where w is a nonstochastic scalar weight, with $0 \leq w \leq 1$.

Weibing Zuo (12) proposed a new weighted stochastic restricted Liu estimator (WSLE) as follows:

$$\underline{\hat{\beta}}_{wsle}(d) = \left(F_d^{-1}S + wH'V^{-1}H\right)^{-1}(X'\underline{Y} + wH'V^{-1}\underline{h}) \quad \ldots(9)$$

Additionally, Hu Yang et al. (13) introduced a weighted mixed Liu estimator as follows:

$$\underline{\hat{\beta}}_w(d) = (S + wH'V^{-1}H)^{-1}(F_d\, X'\underline{Y} + wH'V^{-1}\underline{h}) \quad \ldots(10)$$

Nilgun Yildiz (14) introduced the weighted mixed Liu-type estimator (WMLTE) based on the weighted mixed and Liu-type estimator (LTE) in linear regression model as follows:

$$\underline{\hat{\beta}}\;(w,k,d) = \underline{\hat{\beta}}_{LTE}(k,d) + S_k^{-1}\;H'\left(V + HS_k^{-1}H'\right)^{-1}\left(\underline{h} - H\underline{\hat{\beta}}_{LTE}(k,d)\right), \ldots(11)$$

Where $\underline{\hat{B}}_{LTE}(k,d) = (S + kI)^{-1}\left(X'\underline{Y} - d\underline{\hat{\beta}}\right)$ is the Liu- type estimator.

Additionally, Nimet Özbay and Selahattin Kaçıranlar (15) introduced a new two-parameter-weighted mixed estimator (TPWME) by unifying the weighted mixed estimator in (7) and the two-parameter estimator (TPE) of Özkale and Kaçıranlar (16) as follows:

$$\underline{\hat{\beta}}_w(k,d) = \underline{\hat{\beta}}(k,d) + S_k^{-1}\;H'\left(V + HS_k^{-1}H'\right)^{-1}\left(\underline{h} - H\underline{\hat{\beta}}(k,d)\right), \ldots (12)$$

where $\underline{\hat{\beta}}(k,d) = S_k^{-1}\left(X'\underline{Y} + kd\underline{\hat{\beta}}\right)$.

We want to mention here, that there are many authors who are working in biased estimation methods in regression models with prior information or without, see, for example Özkale (17), Huang and Yang (18) , Kristofer M., Kibria G. B.M. and Shukur G. (19) and Özbay and Kaçıranlar (20).

As it can be observed, all the estimators in (5) to (10) are still dealing with the S^{-1} matrix. Therefore, if there is severe multicollinearity, then the estimators will be obtained but with high variance.

For this reason, the goal of this paper is to propose new types of stochastic-restricted Liu estimators in the case of the prior and the sample information being either equally important or not which does not deal with S^{-1} only.

This paper is organized as follows. In Section 2, the statistical model and the new weighted and non-weighted mixed Liu estimators are introduced. Then, in Section 3, the superiority of the proposed estimator compared with some related estimators is given, and we list some lemmas needed for the theoretical discussions. Finally, a numerical example and a simulation study are provided to illustrate some of the theoretical results in Section 5.

The Proposed Estimators

Case of the prior and sample information are equally important

By augmenting model (1) with $d\underline{\hat{\beta}} = I\underline{\beta} + \underline{\acute{\epsilon}}$, we get

$$\begin{pmatrix} \underline{Y} \\ d\underline{\hat{\beta}} \end{pmatrix} = \begin{pmatrix} X \\ I \end{pmatrix}\underline{\beta} + \begin{pmatrix} \underline{\epsilon} \\ \underline{\acute{\epsilon}} \end{pmatrix}, \; \acute{\epsilon} \sim (0, \sigma^2 I) \text{ or}$$

$$\underline{Y}^* = X^*\underline{\beta} + \underline{\epsilon}^* \quad \ldots \quad (13)$$

By using the mixed estimation method suggested by Thiel and Goldberger (7), the estimation (13) subject to (4), gives the following:

$$\begin{pmatrix} \underline{Y}^* \\ \underline{h} \end{pmatrix} = \begin{pmatrix} X^* \\ H \end{pmatrix}\underline{\beta} + \begin{pmatrix} \underline{\epsilon}^* \\ \underline{e} \end{pmatrix} \quad \ldots (14)$$

In another form, (14) can be rewritten as follows:

$$\underline{\tilde{y}} = \tilde{X}\underline{\beta} + \underline{\tilde{\epsilon}}, \quad ... \quad (15)$$

where $E(\underline{\tilde{\epsilon}}) = 0$, $V(\underline{\tilde{\epsilon}}) = \sigma^2 \tilde{V} = \sigma^2 \begin{pmatrix} I & 0 \\ 0 & V \end{pmatrix}$.

Model (15) combines the prior information in (5) and the sample information in (13). Since V is positive definite (pd), $\tilde{V}$ will be positive definite; therefore, there exists a non-singular symmetric matrix W, such that $\tilde{V} = W'W$ (21). By premultiplying both sides of (15) by W^{-1}, we get the following:

$$\underline{Y}'' = X''\underline{B} + \underline{\epsilon}'', ... \quad (16)$$

where $\underline{Y}'' = W^{-1}\underline{\tilde{Y}}$, $X'' = W^{-1}\tilde{X}$, and $\underline{\epsilon}'' = W^{-1}\tilde{\epsilon}$. From (16), the variance of $\underline{\epsilon}''$ will be equal to $\sigma^2 I$ (i.e., V($\underline{\epsilon}''$)= $\sigma^2 I$). This means that the errors $\underline{\epsilon}''$ are uncorrelated and that (16) represents the classic linear model. Therefore, by fitting model (16) using the least square method, we obtain the new proposed estimator as follows:

$$\underline{\hat{\beta}}_m(d) = (S + I + H'V^{-1}H)^{-1}(X'\underline{Y} + d\underline{\hat{\beta}} + H'V^{-1}\underline{h}) \quad ...(17)$$

The following fact gives another form of $\underline{\hat{\beta}}_m(d)$ in (17).

Lemma 1 (21): Suppose that the square matrices $E: n \times n$ and $C: n \times n$ are not singular, and let $B: p \times n$ and $D: n \times p$ be any two matrices. Then,

$$(E + BCD)^{-1} = E^{-1} - E^{-1}B(C^{-1} + DE^{-1}B)^{-1}DE^{-1}$$

If E= $S + I$, B=H', C =V^{-1} and D =H, then by Lemma 1, we can rewrite (17) as follows:

$$\underline{\hat{\beta}}_m(d) = \underline{\hat{\beta}}(d) + S^{-1}(I)H'(V + HS^{-1}(I)H')^{-1}(\underline{h} - H\underline{\hat{\beta}}(d)), ... (18)$$

where $S^{-1}(I) = (S + I)^{-1}$. We call this estimator the stochastic restricted Liu-type estimator (SRLTE).

Case of the prior and sample information are not equally important

As mentioned in section (2-1), the proposed estimator is as follows:

$$\underline{\hat{\beta}}_{wm}(d) = (S + I + wH'V^{-1}H)^{-1}(X'\underline{Y} + d\underline{\hat{\beta}} + wH'V^{-1}\underline{h})$$

$$= (S^{-1}(I) - wS^{-1}(I)H'(V + wHS^{-1}(I)H')^{-1}HS^{-1}(I))(G_d X'\underline{Y} + wH'V^{-1}\underline{h})$$

$$= S^{-1}(I)G_d X'\underline{Y} + wS^{-1}(I)H'(V + wHS^{-1}(I)H')^{-1}(\underline{h} - HS^{-1}(I)G_d X'\underline{Y})$$

$$= \underline{\hat{\beta}}(d) + wS^{-1}(I)H'(V + wHS^{-1}(I)H')^{-1}\left(\underline{h} - H\underline{\hat{\beta}}(d)\right), \quad ... (19)$$

where $G_d = (I + dS^{-1})$. It is called the weighted-mixed Liu-type estimator (WMLTE).

In fact, $\underline{\hat{\beta}}_{wm}(d)$ is a general estimator, which includes the OLS, LE and SRLTE estimators as special cases. This estimator is as follows.

If $w=0$ and $d=1$, then

$$\underline{\hat{\beta}}_{0m}(1) = \underline{\hat{\beta}}$$

If $w=0$ or $H=0_{q \times p}$, then

$$\underline{\hat{\beta}}_{0m}(d) = \underline{\hat{\beta}}_{wm}(d) = \underline{\hat{B}}_{LE}(d)$$

If $w=1$, then

$$\underline{\hat{\beta}}_{1m}(d) = \underline{\hat{\beta}}_m(d)$$

The properties of the proposed estimators

The properties of the proposed estimator in (19) will be obtained, and then the estimator is generalized using the proposed estimator in (17), by setting $w=1$.

It is well known that the performance of any estimator $\hat{\beta}*$ for β depends on its properties. Therefore, it is necessary to study the properties of the proposed estimator as well as those of other estimators.

The expected value and the variance - covariance matrices of the WME and WSLE estimators are given as follows:

$$E(\underline{\hat{\beta}}_{wm}) = \underline{\beta}, \quad ...(20)$$

$$Cov\left(\underline{\hat{\beta}}_{wm}\right) = \sigma^2 A^*(S + w^2 H'V^{-1}H)A^*, ... (21)$$

where $A^* = (S + wH'V^{-1}H)^{-1}$. Furthermore,

$$E\left(\underline{\hat{\beta}}_{wsle}(d)\right) = \underline{\beta} + A^*(F_d - I)S\underline{\beta} \quad ... \quad (22)$$

$$Cov\left(\underline{\hat{\beta}}_{wsle}(d)\right) = \sigma^2 A^*(F_d SF_d + w^2 H'V^{-1}H)A^* \quad ...(23)$$

Additionally, the expected value and the variance - covariance matrix of the WMLTE are as follows:

$$E\left(\underline{\hat{\beta}}_{wm}(d)\right) = \underline{\beta} + (d-1)A\underline{\beta} \quad ...(24)$$

$$Cov\left(\underline{\hat{\beta}}_{wm}(d)\right) = \sigma^2 A(S + dS^{-1} + w^2 H'V^{-1}H)A \quad ... (25)$$

where $A = (S + I + wH'V^{-1}H)^{-1}$.

If we are dealing with the biased estimators, the mean squared error matrix is the best criterion that can provide good information about the performance of an estimator. This matrix can describe the variance – covariance matrix and the biased vector of an estimator simultaneously as follows:

$$MSE(\underline{\hat{\beta}}^*) = Cov(\underline{\hat{\beta}}^*) + Biased(\underline{\hat{\beta}}^*)\, Biased(\underline{\hat{\beta}}^*)'$$

Therefore,

$$MSE\left(\underline{\hat{\beta}}_{wm}\right) = \sigma^2 A^*(S + w^2 H'V^{-1}H)A^* \quad ...(26)$$

$$MSE\left(\underline{\hat{\beta}}_{wm}(d)\right) = \sigma^2 A(S + dS^{-1} + w^2 H'V^{-1}H)A + B_1 B_1', \quad ... \quad (27)$$

where $B_1 = (d-1)A\underline{\beta}$.

$$MSE\left(\underline{\hat{\beta}}_{wsle}(d)\right) = \sigma^2 A^*(F_d SF_d + w^2 H'V^{-1}H)A^* + B_2 B_2', \quad ... (28)$$

where $B_2 = A^*(F_d - I)S\underline{\beta}$.

Superiority of the Proposed Estimator

In this section, the superiority of $\underline{\hat{\beta}}_{wm}$ (d) will be studied to the other estimators by using the mean squared error matrix. Before that, to clarify the discussion, a definition and some lemmas are listed.

Definition: (22):

Let A: $n \times n$ and B: $n \times n$ be any two matrices. Then, the roots $\lambda_i = \lambda_i^B(A)$ of the equation $|A - \lambda B| = 0$ are called the eigenvalues of A in the metric B.

It is clear from the above definition that the roots of $\lambda_i^B(A)$ are the usual eigenvalues of the matrix $B^{-\frac{1}{2}}AB^{-\frac{1}{2}}$.

Lemma 2 (22)**:**

Let B be a positive definite matrix and A be a positive semi definite matrix denoted by Λ= diag($\lambda_i^B(A)$), which is the diagonal matrix of the eigenvalues of A in the metric B. Then, there exists a nonsingular matrix Q, such that B=Q'Q and A=Q'ΛQ'.

Lemma 3 (21)**:**

Let $\hat{\beta}_1$ and $\hat{\beta}_2$ be two estimators, and let D= Cov($\hat{\beta}_1$) - Cov($\hat{\beta}_2$) be a positive definite matrix. Then, $\Delta(\hat{\beta}_1, \hat{\beta}_2) = MSE(\hat{\beta}_1) - MSE(\hat{\beta}_2)$ is a positive definite if and only if

$$d'_2\ (D + d'_1 d_1)^{-1} d_2 < 1.$$

Lemma 4 (23)

Let M and N be two $n \times n$ matrices, such that M is a positive definite and N is a non-negative definite. Then, M-N is a positive definite if and only if $\lambda_1(NM^{-1}) < 1$.

Superiority of $\underline{\hat{\beta}}_{wm}$ (d) compared to $\underline{\hat{\beta}}_{wm}$

Let $D_1 = \mathrm{Cov}\left(\underline{\hat{\beta}}_{wm}\right) - \mathrm{Cov}\left(\underline{\hat{\beta}}_{wm}\ (\mathrm{d})\right)$

$= \sigma^2 A^*(S + w^2 R'V^{-1}R)\, A^* - \sigma^2 A(S + dS^{-1} + w^2 H'V^{-1}H)\mathrm{A}$.

By lemma 2, let Q be a nonsingular matrix, such that

$\sigma^2 A^*(\Lambda + w^2 R'V^{-1}R)\, A^* = Q'Q$ and $\sigma^2 A(S + dS^{-1} + w^2 H'V^{-1}H)\mathrm{A} = Q'\Lambda Q$.

Therefore, $D_1 = Q'Q - Q'\Lambda Q = Q'(I - \Lambda)Q$.

Let $x'D_1x = x'Q'(I - \Lambda)Qx = y'(I - \Lambda)y$

$= \sum\left(1 - \lambda_i^{\sigma^2 A^*(\Lambda + w^2 H'V^{-1}H)\, A^*}(\sigma^2 A(S + dS^{-1} + w^2 H'V^{-1}H)\mathrm{A})\right) y_i^2$,

where $y = Qx$.

It is clear that $x'D_1x$ is p.d. if and only if $\lambda_i^{\sigma^2 A^*(\Lambda + w^2 H'V^{-1}H)\, A^*}(\sigma^2 A(S + dS^{-1} + w^2 H'V^{-1}H)\mathrm{A}) < 1$.

Therefore, by lemma 3, the following result can be stated.

Theorem 1: If $\lambda_i^{\sigma^2 A^*(\Lambda + w^2 H'V^{-1}H)\, A^*}(\sigma^2 A(S + dS^{-1} + w^2 H'V^{-1}H)\mathrm{A}) < 1$ for some i, then $\Delta_1 = MSE\left(\underline{\hat{\beta}}_{wm}\right) - MSE\left(\underline{\hat{\beta}}_{wm}\ (\mathrm{d})\right)$ is a positive definite if and only if $B'_1\ D_1^{-1} B_1 < 1$.

Superiority of $\underline{\hat{\beta}}_{wm}$ (d) compared to $\underline{\hat{\beta}}_{wsle}(d)$

To compare these estimators, we similarly consider the difference $MSE\left(\underline{\hat{\beta}}_{wsle}(d)\right) - MSE\left(\underline{\hat{\beta}}_{wm}\ (d)\right)$ as follows:

$$\mathrm{MSE}\left(\underline{\hat{\beta}}_{wsle}(d)\right) - \mathrm{MSE}\left(\underline{\hat{\beta}}_{wm}\ (d)\right) = \sigma^2 \mathrm{D}_2 + \mathrm{B}_2\mathrm{B}_2' - \mathrm{B}_1\mathrm{B}_1',$$

where $\mathrm{D}_2 = \sigma^2 \mathrm{A}^*\left(\mathrm{F_d\ SF_d} + w^2 H'V^{-1}H\right)\mathrm{A}^* - \sigma^2 \mathrm{A}(\mathrm{S} + dS^{-1} + w^2 H'V^{-1}H)\mathrm{A}$.

According to lemma 4, D_2 is a positive definite if and only if

$$\lambda_1\left(\sigma^2 \mathrm{A}(\mathrm{S} + dS^{-1} + w^2 H'V^{-1}H)\mathrm{A}(\sigma^2 \mathrm{A}^*\left(\mathrm{F_d\ SF_d} + w^2 H'V^{-1}H\right)\mathrm{A}^*)^{-1}\right) < 1.$$

Now, the following result can be stated.

Theorem 2:

If $\lambda_1\left(\sigma^2 \mathrm{A}(\mathrm{S} + dS^{-1} + \mathrm{w}^2\mathrm{R'V^{-1}R})\mathrm{A}(\sigma^2 \mathrm{A}^*\left(\mathrm{F_d\ SF_d} + \mathrm{w}^2\mathrm{R'V^{-1}R}\right)\mathrm{A}^*)^{-1}\right) < 1$, then the weighted-mixed Liu-type estimator $\underline{\hat{\beta}}_{\mathrm{wm}}\,(d)$ is superior to the weighted-stochastic-restricted Liu-type estimator $\underline{\hat{\beta}}_{\mathrm{wsle}}(d)$ in the mean squared error matrix if and only if

$$B'_1\ (D + B'_2 B_2)^{-1} B_1 < 1.$$

Optimal biased parameter d for $\underline{\hat{\beta}}_{\mathrm{wm}}$ (d)

The least mean squared error of $\underline{\hat{\beta}}_{\mathrm{wm}}\,(d)$ can be obtained by finding the optimal value of the biased parameter d. Therefore, we have to find the d that achieves the desired performance.

Model (1) can be written in the canonical form as follows:

$$\underline{Y} = Z\underline{\alpha} + \underline{\epsilon},$$

where $Z = XP$, $\underline{\alpha} = P'\underline{\beta}$ and P is a $p \times p$ orthogonal matrix, such that $P'X'XP = \Lambda$. Λ is a $p \times p$ diagonal matrix and its elements $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the eigenvalues of $Z'Z$, such that $\lambda_1 > \lambda_2 > \cdots > \lambda_p$ and $P'H'V^{-1}HP = \Gamma = diag\{v_1, \ldots, v_p\}$.

Therefore, the MSE of the proposed estimator is as follows:

$MSE\left(\underline{\hat{\alpha}}_{wm}(\text{d})\right) = \sigma^2(\Lambda + I + w\Gamma)^{-1}(\Lambda + d\Lambda^{-1} + w^2\Gamma)(\Lambda + I + w\Gamma)^{-1} + (d-1)^2(\Lambda + I + w\Gamma)^{-1}\underline{\alpha}\underline{\alpha}'(\Lambda + I + w\Gamma)^{-1}$.

To find the optimal value of the biased parameter d that minimizes $MSE\left(\underline{\hat{\alpha}}_{wm}(d)\right)$, let w be fixed, minimize the trace of the $MSE\left(\underline{\hat{\alpha}}_{wm}(\text{d})\right)$ as a function and calculate the derivative with respect to d as follows:

$$mse = tr\{MSE\left(\underline{\hat{\alpha}}_{wm}(d)\right)\} = \sum_{i=1}^{p} \frac{\sigma^2(\lambda_i + d\lambda_i^{-1} + w^2 v_i) + (d-1)^2\alpha_i^2}{(\lambda_i + 1 + wv_i)^2}$$

Thus,

$$\frac{\partial\, mse}{\partial\, d} = \sum_{i=1}^{p} \frac{\sigma^2\lambda_i^{-1} + 2(d-1)\ \alpha_i^2}{(\lambda_i + 1 + wv_i)^2} = 0$$

After some simplifications, the optimal d will be given as follows:

$$d = 1 - \frac{\sigma^2 \sum_{i=1}^{p} \lambda_i^{-1}}{2\sum_{i=1}^{p} \alpha_i^2}. \quad \dots \quad (29)$$

The optimal value of d in (26) depends on two unknown parameters, σ^2 and α_i^2. Therefore, these parameters are replaced with their unbiased estimators $\hat{\sigma}^2$ and $\hat{\alpha}_i^2$ to get the following (see (24)):

$$\hat{d} = 1 - \frac{\hat{\sigma}^2 \sum_{i=1}^{p} \lambda_i^{-1}}{2\sum_{i=1}^{p} \hat{\alpha}_i^2}. \quad (30)$$

Numerical Example and Simulation Study

In this section, the performance of the new estimator is explained compared to the other estimators (WME and WSLE) using the scalar mean squared error (mse). We use the dataset on Portland Cement originally attributed to Woods et al. (25), which several researchers used in their studies, including Hu Yang and Jianwen Xu (9), Hu Yang et al. (26). Our computations were all performed using Matlab R2010b.

The following stochastic linear restriction is considered to improve estimator:

$\underline{h} = H\underline{\beta} +$e, H=(1, 1,1,0) and $e \sim (0, \sigma^2{}_{OLS})$ (see (27)).

- **The performance of WMLTE with respect to *w*.**

From Tables 1,2,3 for different values of *w*, in most cases (when d < 0.35), the WMLTE is better than the other estimators, which can be observed in Figures 1-5.

- **The performance of WMLTE with respect to *d***

The minimum mse value for the WMLTE is when *d*=0.01, but this is not the case for the WML and WSLE. Furthermore, when *d*=0.99, the mse of WMLTE is its maximum value, and the other estimators are superior. These results are clear by looking at Tables 1-3. Therefore, we can say that there is a relation between the performance of the WMLTE and d, where if *d* increases, then the mse of WMLTE will decrease and vice versa. Figures ,2,3,4,5 show the performance of the WMLTE with respect to d for different values of *w*.

Table 1. The scalar mean squared error for the WME, WMLTE and WSLE when w=0.05 and w= 0.1

w=0.1				w=0.05			
mse (WSLE)	mse (WMLTE)	mse (WME)	d	mse (WSLE)	mse (WMLTE)	mse (WME)	d
1.077	0.254	0.7489	0.01	0.8271	0.2585	0.805	0.01
0.9216	0.313	0.7489	0.09	0.7089	0.317	0.805	0.09
0.8199	0.3593	0.7489	0.15	0.6337	0.363	0.805	0.15
0.7449	0.3992	0.7489	0.2	0.5798	0.4026	0.805	0.2
0.6666	0.4488	0.7489	0.26	0.5257	0.4519	0.805	0.26
0.6215	0.4828	0.7489	0.3	0.496	0.4858	0.805	0.3
0.573	0.5264	0.7489	0.35	0.4661	0.5292	0.805	0.35
0.5407	0.5622	0.7489	0.39	0.448	0.5648	0.805	0.39
0.5082	0.608	0.7489	0.44	0.4325	0.6105	0.805	0.44
0.4808	0.6647	0.7489	0.5	0.4244	0.667	0.805	0.5
0.4663	0.7231	0.7489	0.56	0.4279	0.7253	0.805	0.56
0.4644	0.7832	0.7489	0.62	0.4428	0.7854	0.805	0.62
0.4727	0.8347	0.7489	0.67	0.4641	0.8368	0.805	0.67
0.4942	0.8981	0.7489	0.73	0.5002	0.9002	0.805	0.73
0.5285	0.9633	0.7489	0.79	0.5478	0.9654	0.805	0.79
0.5584	1.0077	0.7489	0.83	0.5859	1.0098	0.805	0.83
0.6038	1.0644	0.7489	0.88	0.6408	1.0665	0.805	0.88
0.6465	1.1106	0.7489	0.92	0.6904	1.1128	0.805	0.92
0.7349	1.1933	0.7489	0.99	0.7896	1.1957	0.805	0.99

Table 2. The scalar mean squared error for the WME, WMLTE and WSLE when w=0.35 and w= 0.75

w=0.75				w=0.35			
mse(WSLE)	mse (WMLTE)	mse (WME)	d	mse (WSLE)	mse (WMLTE)	mse (WME)	d
1.5166	0.2365	0.716	0.01	1.4124	0.2426	0.7181	0.01
1.2966	0.2975	0.716	0.09	1.2076	0.3029	0.7181	0.09
1.1496	0.3451	0.716	0.15	1.0712	0.35	0.7181	0.15
1.0387	0.3861	0.716	0.2	0.9688	0.3906	0.7181	0.2
0.9198	0.4368	0.716	0.26	0.8593	0.4409	0.7181	0.26
0.8491	0.4714	0.716	0.3	0.7945	0.4753	0.7181	0.3
0.7703	0.5158	0.716	0.35	0.7227	0.5194	0.7181	0.35
0.7149	0.5521	0.716	0.39	0.6726	0.5556	0.7181	0.39
0.6553	0.5986	0.716	0.44	0.6192	0.6018	0.7181	0.44
0.5978	0.6558	0.716	0.5	0.5686	0.6588	0.7181	0.5
0.5557	0.7147	0.716	0.56	0.5326	0.7175	0.7181	0.56
0.5289	0.7752	0.716	0.62	0.5114	0.7779	0.7181	0.62
0.5184	0.8268	0.716	0.67	0.5049	0.8295	0.7181	0.67
0.5197	0.8903	0.716	0.73	0.5106	0.8929	0.7181	0.73
0.5365	0.9555	0.716	0.79	0.531	0.9581	0.7181	0.79
0.5562	0.9998	0.716	0.83	0.5528	1.0025	0.7181	0.83
0.5904	1.0563	0.716	0.88	0.5891	1.059	0.7181	0.88
0.6254	1.1023	0.716	0.92	0.6256	1.105	0.7181	0.92
0.7032	1.1845	0.716	0.99	0.7052	1.1875	0.7181	0.99

Table 3. The scalar mean squared error for the WME, WMLTE and WSLE when w=0.95

w=0.95			
mse(WSLE)	mse(WMLTE)	mse(WME)	d
1.5376	0.2351	0.7159	0.01
1.3145	0.2963	0.7159	0.09
1.1654	0.344	0.7159	0.15
1.0529	0.3851	0.7159	0.2
0.9321	0.4358	0.7159	0.26
0.8601	0.4706	0.7159	0.3
0.7799	0.515	0.7159	0.35
0.7235	0.5514	0.7159	0.39
0.6626	0.5979	0.7159	0.44
0.6038	0.6551	0.7159	0.5
0.5605	0.714	0.7159	0.56

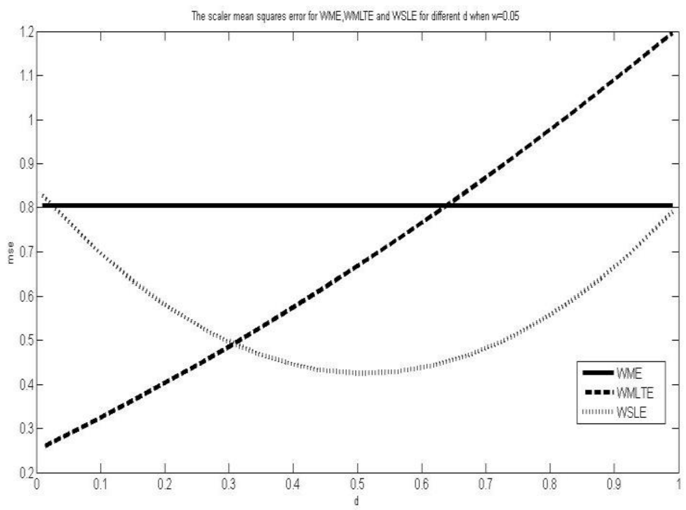

Figure 1. The mean squared error of the WME, WMLTE and WSLE for *w=0.05*

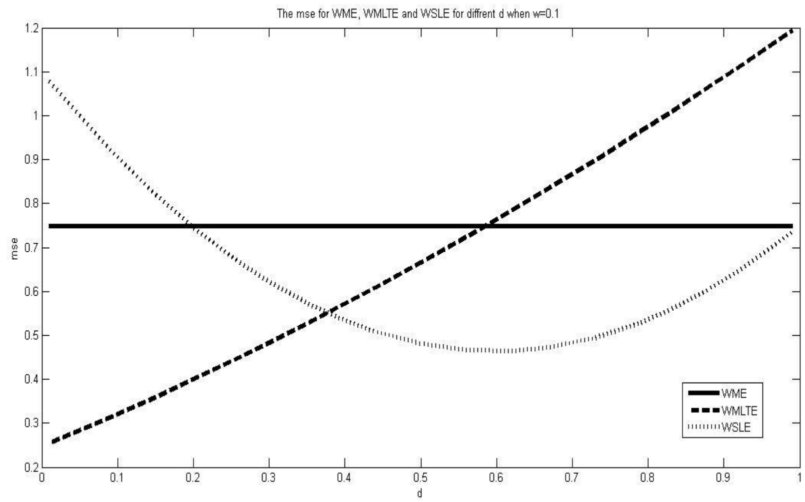

Figure 2. The mean squares error of the WME, WMLTE and WSLE for w=0.1

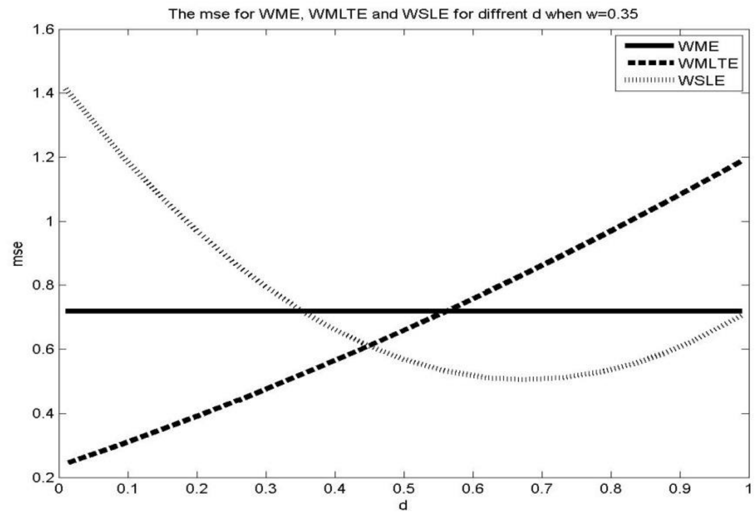

Figure 3. The mean squared error of the WME, WMLTE and WSLE for w=0.35

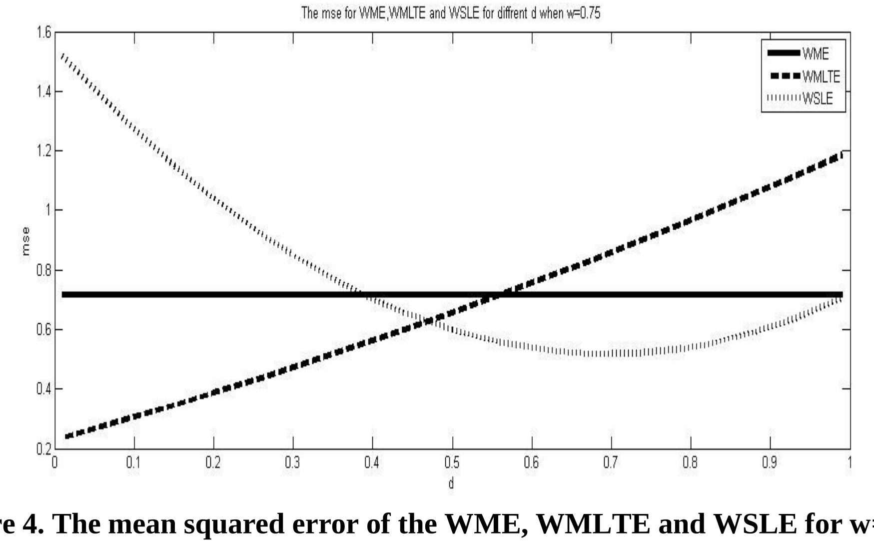

Figure 4. The mean squared error of the WME, WMLTE and WSLE for w=0.75

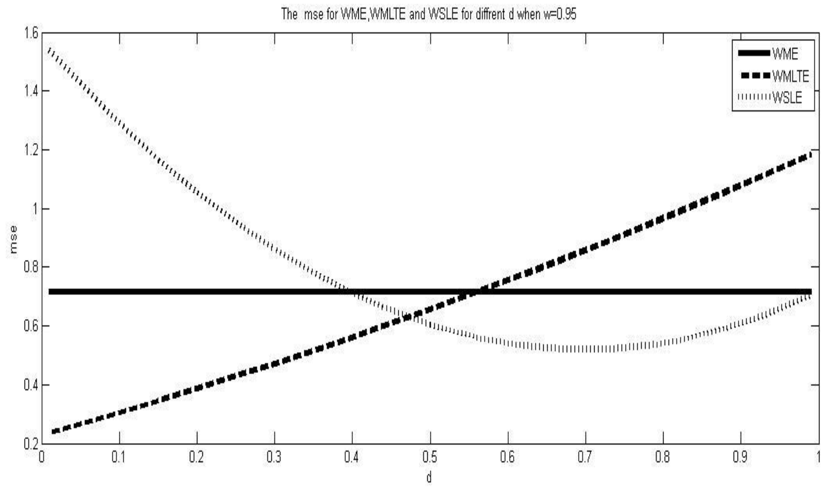

Figure 5. The mean squared error of the WME, WMLTE and WSLE for w=0.95

For further explanation regarding the behaviour of the new estimator, a Monte Carlo simulation experiment was performed. Following Kibria and Banik (24) and Hua Huang et al. (18) to achieve various degrees of collinearity, the explanatory variables are generated by using the following equation.

$$x_{ij} = (1-\rho^2)^{\frac{1}{2}} z_{ij} + \rho z_{i5}, \quad i = 1,2,\dots,n, \quad j = 1,\dots,p,$$

where z_{ij} are independent standard normal pseudo-random numbers, p=4 is the number of the explanatory variables, $n=100$ and 500, and ρ is specified so that the correlation between any two explanatory variables is given by ρ^2. The observations of the dependent variable are then generated by the following:

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

where ϵ_i are independent normal pseudorandom numbers, with a mean of zero and a variance σ^2. In this study, we choose $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)' = (1,2,3,4)'$, $\sigma^2 = 1$ and $\rho = 0.85, 0.95$ and 0.99. The experiment is replicated 2000 times by generating new error terms. The MSE for the estimators is calculated as follows: $\beta^* = \frac{1}{2000}\sum_{i=1}^{2000}\left(\beta^*_{ij} - \beta_i\right)^2$, where β^*_{ij} is the estimator of the ith parameter in the jth replication and β_i is the true parameter value.

Table 4.The estimated MSE for the WME, WMLTE and WSLE when n=100

	$\rho = 0.85$			$\rho = 0.95$			$\rho = 0.99$		
w	WME	WSLE	WMLTE	WME	WSLE	WMLTE	WME	WSLE	WMLTE
0.05	0.21117	0.21169	0.21032	0.23003	0.23089	0.22515	0.2522	0.25474	0.24033
0.1	0.21283	0.21334	0.20963	0.23378	0.23464	0.22468	0.25458	0.25746	0.2398
0.35	0.2183	0.2188	0.20906	0.23932	0.24028	0.22524	0.25664	0.2599	0.24057
0.75	0.22133	0.22184	0.21024	0.24113	0.24213	0.22737	0.25713	0.26048	0.243
0.95	0.22205	0.22257	0.21083	0.24149	0.24251	0.22821	0.25722	0.26059	0.24392

Table 5. The estimated MSE for the WME, WMLTE and WSLE when n=500

	$\rho = 0.85$			$\rho = 0.95$			$\rho = 0.99$		
w	WME	WSLE	WMLTE	WME	WSLE	WMLTE	WME	WSLE	WMLTE
0.05	0.21485	0.21495	0.21256	0.2308	0.231	0.22385	0.2453	0.24619	0.23033
0.1	0.21725	0.21735	0.21209	0.23481	0.23502	0.2235	0.24752	0.24849	0.23012
0.35	0.22353	0.22363	0.21226	0.24036	0.24058	0.22445	0.24942	0.25046	0.23157
0.75	0.2266	0.2267	0.21393	0.24208	0.24232	0.22683	0.24987	0.25092	0.23424
0.95	0.2273	0.22741	0.21465	0.24243	0.24267	0.22773	0.24995	0.25101	0.2352

Table 4 and Table 5 show that the WMLTE estimator is better than the other estimators for different values of correlation and for both cases ($n=100$ and $n=500$). This result supports the goal of this article for finding or improving an estimator that is more accurate compared to other estimators. It is clear that the new estimator is meaningful in practice.

Conclusion:

In this paper, a new version of the weighted-mixed Liu-type estimator is introduced for the vector of parameters in a linear regression model by unifying the sample and the prior information in the case that they either are equally or are not equally important. Furthermore, the new estimator is superior to the weighted-mixed estimator and the weighted-stochastic-restricted Liu-type estimator in the mean squared error matrix under certain conditions. The optimal value of the biased parameter for the proposed estimator are obtained. Finally, a numerical example and a simulation study are given for the comparison of the new estimator with other estimators in this study.

Acknowledgements:

The author is thankful to the anonymous referees and the Editor in Chief for their valuable comments and suggestions that helped to improve the quality of the article.

Conflicts of Interest: None.

References

1. Neter W,Kutner. Applied Linear Regression Models, 2nd edition. Irwin, Homewood IL. 1989.
2. Salmerón R, García C B , García J. Variance Inflation Factor and condition Number in multiple linear regression. J Statist Comp Simu. 2018; 88(12):2365-2384.
3. Hoerl AE, Kennard RW. Ridge regression: biased estimation for non-orthogonal problems. Technometrics . 1970 Feb; 12:55–67.
4. Liu K. A new class of biased estimate in linear regression. Commun Statist Theor Meth. 1993; 22:393–402.
5. Fela O, Selahattin k. Evaluation of the Predictive Performance of the Liu Estimator. Commun Statist Theor Meth. 2015 May 29; 44: 1981–1993.
6. Sivarajah A , Pushpakanthie W. Modified Almost Unbiased Liu Estimator in Linear Regression Model. Commun Math Statis. 2017; 5(3): 261–276.
7. Theil H, Goldberger A S. On pure and mixed statistical estimation in economics. Int Econ Rev. 1961; 2:65–78.
8. Theil H. On the use of incomplete prior information in regression analysis. J Amer Statist Assoc. 1963;58:401–414.
9. Hu Yang , Jianwen Xu. An alternative stochastic restricted Liu estimator in linear Regression. Stat Papers. 2009 Jun; 50:639–647.
10. Nilgün Y. **A** New Stochastic Restricted Liu-Type Estimator in Linear Regression Model. Commun Statist Simu Comp. 2019; 48(1): 91-108.
11. Schaffrin B, Toutenburg H. Weighted mixed regression. Zeitschrift fur Angewandte Mathematik und Mechanik. 1990; 70:735–738.
12. Weibing Zuo. A New Stochastic Restricted Liu Estimator in Weighted Mixed Regression. Sec Int Symp on Comp Intelli and Desi . 2009; 245-248.
13. Hu Yang , Xinfeng Chang , Deqiang Liu. Improvement of the Liu Estimator in Weighted Mixed Regression. Commun Statist Theor Meth. 2009 ;38: 285–292.
14. Nilgün Y. On the Weighted Mixed Liu-Type Estimator Under Unbiased Stochastic Restrictions. Commun Statist Simu Comp. 2017 Apr 21; 46(9): 7238-7248.
15. Nimet Ö , Selahattin K. Estimation in a linear regression model with stochastic linear restrictions: a new two-parameter-weighted mixed estimator. J statist Comp and simu 2018 Mar 05; 88(9): 1669–1683.
16. Özkale MR, Kaçıranlar S. The restricted and unrestricted two parameter estimators. Commun Statist Theor Meth. 2007; 36(15):2707–2725.
17. Özkale MR. The relative efficiency of the restricted estimators in linear regression models. J Appl Stat. 2014; 41(5): 998–1027.
18. Huang J,Yang H. A two-parameter estimator in the negative binomial regression model. J Statist Comput Simul. 2014; 84(1): 124–134.
19. Kristofer M, Kibria G BM, Shukur G. A restricted Liu estimator for binary regression models and its application to an applied demand system. J App Statis. 2016 ;43(6): 1119–1127.
20. Özbay N, Kaçıranlar S. The Almon two parameter estimator for the distributed lag models. J Statist Comp Simu. 2017; 87(4): 834–843.
21. Rao C R, Toutenburg H, Shalabh, Heumann C. Linear Models and Generalizations— Least Squares and Alternatives. Berlin: Springer; 2008.
22. Rao C R, Toutenburg H. Linear models: least squares and alternatives. Springer, New York; 1995.
23. Wang S G, Wu M X, Jia Z Z. Matrix Inequalities. 2nd ed. Beijing: Chinese Science Press;2006.
24. Kibria G BM, Banik S. Some Ridge Regression Estimators and Their Performances. J Mod App Statis Meth. 2016; 15(1): 206-238.
25. Woods H, Steinour H H, Starke H R. Effect of composition of Portland cement on heat evolved during hardening. Indust Eng Chem. 1932; 24:1207–1241
26. Hua Huang, Jibo Wu and Wende Yi On the restricted almost unbiased two parameter estimator in linear regression model. Commun. Statist. Theor. Meth. 2017; 46(4): 1668-1678.
27. Kristofer M, Kibria G BM, Shukur G. Performance of Some Weighted Liu Estimators for Logit Regression Model: An Application to Swedish Accident Data. Commun Statist Theor Meth. 2015; 44: 363–375.

Open Access Baghdad Science Journal P-ISSN: 2078-8665
2020, 17(1) Supplement (March):361-370 E-ISSN: 2411-7986

إصدارات جديدة لمقدِّر ليو- تايب في نموذج الانحدار المختلط المرجَّح وغير المرجَّح

مصطفى اسماعيل نايف الهيتي

قسم الرياضيات، كلية التربية للعلوم الصرفة ، جامعة الانبار، الانبار، العراق.

الخلاصة:

هذا البحث يصف ويقترح مقدرين جديدين بالاعتماد على معلومات العينة وكذلك المعلومات الاولية في حالة كانت هذه المعلومات متساوية من ناحية الاهمية في بناء النموذج ام غير متساوية. لقد كانت المعلومات الاولية موصوفة كقيود تصادفية خطية. سنقوم بدراسة خواص واداء هذه المقدرات المقترحة مقارنة مع مقدرات اخرى معروفة وذلك من خلال استخدام متوسط مربعات الخطأ كمعيار لجودة التقدير. مثال عددي وكذلك دراسة محاكاة تم اقتراحهما لتوضيح سلوك واداء المقدرات المقترحة في هذا البحث.

الكلمات المفتاحية : مقدر ليو، مصفوفة متوسط مربعات الخطأ، مقدر ليو المقيد التصادفي، مقدر ليو المختلط المرجح، تصنيف موضوع **الرياضيات: اولي 62J05; ثانوي: 62J07**