**Iraqi Journal of Statistical Science (20) 2011**
**The Fourth Scientific Conference of the College of Computer Science & Mathematics**
pp [26-34]

# A new almost unbiased estimator in stochastic linear restriction model

Mustafa Ismaeel Naif*

**Abstract**

In this paper, a new almost unbiased estimator is proposed under stochastic linear restrictions model as alternative to mixed estimator. The performance of the proposed estimator compared to mixed estimator is examined using the matrix mean squared errors.

**Keywords: Liu-type estimator; mixed estimator; stochastic linear restrictions.**

.

## 1 Introduction

Consider the following multiple linear regression model

$$Y = X\beta + \epsilon, \qquad (1)$$

- Lecturer\ Anbar University\ Email: alheety@yahoo.com

Received:1/10 /2011 ____________________Accepted: 21 /12 / 2011

where $\mathbf{Y}$ is an n $\times 1$ vector of observations, $\mathbf{X}$ is an n $\times$p matrix , $\boldsymbol{\beta}$ is a p $\times 1$ vector of unknown parameters, and $\epsilon$ is an n $\times 1$ vector of non observable errors with $\mathrm{E}(\epsilon) = 0$ and $Cov(\epsilon) = \sigma^2 I_n$.

The most common method used for estimating the regression coefficients in (1) is the ordinary least squares (OLS) method which is defined as

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

Both the OLS estimator and its covariance matrix heavily depend on the characteristics of the X'X matrix. If X'X is ill-conditioned, i.e. the column vectors of X are linearly dependent; the OLS estimators are sensitive to a number of errors. For example, some of the regression coefficients may be statistically insignificant can't or have the wrong sign, and they may result in wide confidence intervals for individual parameters. With ill-conditioned X'X matrix, it is difficult to make valid statistical inferences about the regression parameters. One of the most popular estimator dealing with multicollinearity is the ordinary ridge regression (ORR) estimator proposed by Hoerl and Kennard (1970a,b) and defined as

$$\tilde{\beta}_k = (X'X + kI)^{-1}X'Y = [I + k(X'X)^{-1}]^{-1}\hat{\beta} \quad ,k \geq 0 \quad (3)$$

Both of the Liu estimator (LE) and the generalized Liu estimator (GLE) are defined (see Kaciranlar et al., 1999) as follows:

$$\tilde{\beta}_d = (X'X + I)^{-1}(X'Y + d\,\hat{\beta}) \quad 0 < d < 1 \quad (4)$$

The advantage of the LE over the ORR is that the LE is a linear function of d, so it is easy to choose d than to choose k in the ORR estimator.

Since X'X is symmetric, there exists a $p \times p$ orthogonal matrix P such that P'X'XP = $\Lambda$ , $\Lambda$ is a $p \times p$ diagonal matrix, where each element of it is the eigenvalue of X'X. So, model (1) can be written in the canonical form as:

$$Y = Z\alpha + \epsilon, \qquad (6)$$

where $Z = XP$ and $\alpha = P'\beta$. The OLS , ORR and Liu estimators for (6) are respectively:

$$\hat{\alpha} = \Lambda^{-1} Z'Y, \qquad (7)$$

$$\hat{\alpha}_k = (\Lambda + kI)^{-1} Z'Y, \qquad (8)$$

and

$$\hat{\alpha}_d = (\Lambda + I)^{-1}(\Lambda + dI)\hat{\alpha} \qquad (9)$$

In order to reduce the cost of the bias in biased estimators with small change in the variance, Singh et al. (1986) introduced the almost unbiased ridge estimator (AURE) as

$$\hat{\alpha}_{AUR} = [I - k^2(\Lambda + kI)^{-2}]\hat{\alpha} \qquad (10)$$

Also, Akdeniz and Kaciranlar (1995) proposed the almost unbiased generalized Liu estimator

$$\hat{\alpha}_{AUL} = [I - (\Lambda + I)^{-2}(1 - d)^2]\hat{\alpha} \qquad (11)$$

Özkale and Kaçiranlar (2007) introduced a new two-parameter estimator by grafting the Contraction estimator into the modified ridge estimator proposed by Swindel (1976). This new two-parameter estimator is a general estimator which includes OLS, ORR, LE and the contraction estimators as special cases. Their estimator is given as follows:

$$\hat{\alpha}_{(k,d)} = (\Lambda + kI)^{-1}(\Lambda + kdI)\hat{\alpha} \qquad (12)$$

In addition to model (6), we suppose that, there are some prior information about $\alpha$ in the form of a set of independent stochastic linear restrictions

$$r = R\alpha + \epsilon^* \tag{13}$$

where $R$ is a q $\times$p non zero matrix with rank $R$= q < p, $r$ is an q $\times$1 known vector which is interpreted as a random variable with $E(r) = R\alpha$ and $\epsilon^*$ is an q $\times$1 vector of disturbances with zero mean and variance-covariance matrix $Cov(\epsilon) = \sigma^2 V$, $V$ is known and positive definite.

By putting the sample and prior information in a common model we get:

$$\begin{pmatrix} Y \\ r \end{pmatrix} = \begin{pmatrix} Z \\ R \end{pmatrix} + \begin{pmatrix} \epsilon \\ \epsilon^* \end{pmatrix}, \tag{14}$$

Where $\mathrm{E}(\epsilon\epsilon^*) = 0$ and $\mathrm{E}(\epsilon\epsilon^*)\begin{pmatrix} \epsilon \\ \epsilon^* \end{pmatrix} = \sigma^2 \begin{pmatrix} \mathrm{I} & 0 \\ 0 & \mathrm{V} \end{pmatrix}$. If we use the least squares method for the model (14), then we will get the mixed estimator (ME) that introduced by Theill and Goldberger (1961). The ME is defined as follows:

$$\hat{\alpha}_{ME} = (\Lambda + \mathrm{R'V^{-1}R})^{-1}(\mathrm{Z'Y} + \mathrm{R'V^{-1}r}) \tag{15}$$

# 2 The Proposed Estimator and its properties

In order to get the proposed estimator, we first should construct the almost unbiased (k-d) class estimator as follows:

We can rewrite $\hat{\alpha}_{(k,d)}$ in the following form:

$$\hat{\alpha}_{(k,d)} = [\mathrm{I} - \mathrm{k}(\Lambda + \mathrm{kI})^{-1}(1 - \mathrm{d})\ ]\hat{\alpha}. \tag{16}$$

By using kadiyala (1984) technique and Ohtani (1986) procedure, we get the almost unbiased (k-d) class estimator (AU(k-d)) as:

$$\hat{\alpha}_{AU(k,d)} = [\mathrm{I} - \mathrm{k}^2(\Lambda + \mathrm{kI})^{-2}(1 - \mathrm{d})^2]\hat{\alpha}, \tag{17}$$

Where k > 0 and $-\infty < d < \infty$.

Now we can introduce the proposed estimator as alternative to mixed estimator by using (AU(k-d)) estimator as follows:

Let us rewrite the ME in (15) in other form

$$\hat{\alpha}_{ME} = (\Lambda + R'V^{-1}R)^{-1}(Z'Y + R'V^{-1}r)$$

$$= \hat{\alpha} + \Lambda^{-1}R'(V + R'\Lambda^{-1}R)^{-1}(r - R\hat{\alpha}). \quad (18)$$

By substituting $\hat{\alpha}$ with $\hat{\alpha}_{AU(k,d)}$ in (18) we get the proposed estimator:

$$\hat{\alpha}_{AUM(k,d)} = \hat{\alpha}_{AU(k,d)} + \Lambda^{-1}R'(V + R'\Lambda^{-1}R)^{-1}\left(r - R\hat{\alpha}_{AU(k,d)}\right).$$

$$= (\Lambda + R'V^{-1}R)^{-1}(MZ'Y + R'V^{-1}r), \quad (19)$$

Where $M = I - k^2(\Lambda + kI)^{-2}(1-d)^2$.

We call the proposed estimator in (19) the almost unbiased mixed (k-d) estimator (AUM(k-d)).

One of the methods used to reduce inflation in the variance of an estimator, especially in the case of overlapping linear between the independent variables, is to use the estimate biased, where, although the capabilities of least squares is biased, but it has a significant inflation variability. Therefore, statisticians researchers use the biased method to reduce the variance where there is some cost in the amount of bias. The idea now, with a biased estimate is to look for the possibility of reducing the bias at the same time to obtain an improvement in variance. For this, the idea of finding the almost unbiased estimators that are given by kadiyala (1984) and Ohtani (1986) appeared. Therefore, the ME estimator in (15) that depends on the OLS as given in (19) will suffer from the problems of the OLS.

To get the properties of the proposed estimator, we have to give some lemmas that help to understand the results.

***Lemma 1***: (**See C.R. Rao et al. 2008**)

Let A: p × p, B:p×n, C: n× n and D:n ×p. If all the inverses exist, then

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

***Lemma 2***: (**See Farebrother 1976**)

Let A be a p.d. matrix , c be an non zero vector and $\theta$ be a positive scaler. Then $\theta$ A-cc' is p.d. if and only if $c'A^{-1}c < \theta$ .

***Lemma 3***: (**See C.R. Rao et al. 2008**)

Let $\hat{\alpha}_j = A_j Y, j = 1,2$ be two linear estimators of $\alpha$. Suppose that D = $Cov(\hat{\alpha}_1) - cov(\hat{\alpha}_2)$ is p.d. then $\Delta =$ MSE($\hat{\alpha}_1$) - MSE($\hat{\alpha}_2$) is n.n.d. if and only if

$$B_2 (D + B_1 B_1')^{-1} B_2 \leq 1,$$

where $B_j$ denotes the bias vector of $\hat{\alpha}_j$.

The proposed estimator is a general estimator which includes OLS, ME and AU(k-d) estimators. The expected value, the variance and the bias of the AUM(k-d) is given as follows:

Let $A = (\Lambda + R'V^{-1}R)^{-1}$. Then

$$E(\hat{\alpha}_{AUM(k,d)}) = A(M - I)\Lambda\alpha + \alpha. \qquad (20)$$

Therefore,

$$B(\hat{\alpha}_{AUM(k,d)}) = A(M - I)\Lambda\alpha. \quad (21)$$

And

$$Var(\hat{\alpha}_{AUM(k,d)}) = \sigma^2 A(M\Lambda M' + R'V^{-1}R)A. \quad (22)$$

The bias and the variance of an estimator $\hat{\alpha}$ are measured simultaneously by the mean squares error matrix (MSE)

$MSE(\hat{\alpha}^*) = Var(\hat{\alpha}^*) + Bias(\hat{\alpha}^*)(Bias(\hat{\alpha}^*))'$. Then

$$MSE(\hat{\alpha}_{AUM(k,d)}) = \sigma^2 A(M\Lambda M' + R'V^{-1}R)A' + A(M - I)\Lambda\alpha\alpha'\Lambda(M - I)A'. \quad (23)$$

## 3 Superiority of the new estimator

In this section we make a comparison between the proposed estimator and the ME estimator using MSE. Let us consider the difference among the estimators:

$$\Delta = MSE(\hat{\alpha}_{ME}) - MSE(\hat{\alpha}_{AUM(k,d)})$$

$$= \sigma^2 A(\Lambda + R'V^{-1}R)A' - \sigma^2 A(M\Lambda M' + R'V^{-1}R)A' - A(M - I)\Lambda\alpha\alpha'\Lambda(M - I)A'$$

$$= \sigma^2 A[\Lambda - M\Lambda M']A' - A(M - I)\Lambda\alpha\alpha'\Lambda(M - I)A' = \sigma^2 A[\Lambda - M\Lambda M']A' - B_1B'_1.$$

$$= \sigma^2 D - B_1B'_1$$

Where $D = A[\Lambda - M\Lambda M']A'$. In this time we want to see under which condition, $\Delta$ will be p.d. If we use Lemma 3 we can see that if D is p.d. then $\Delta$ will be p.d. if and only if $B_1'D^{-1}B_1 < \sigma^2$. That means we are searching for the condition that makes D p.d.

$D = A[\Lambda - M\Lambda M']A' = A\Lambda[I - MM']A'$. Since A and $\Lambda$ are p.d., we are constricting our self to the difference $[I - MM']$.

$[I - MM'] = diag\{g_1, \dots, g_p\}$, where $g_i = 1 - [1 - k^2(\lambda_i + k)^{-2}(1-d)^2]^2$. Therefore $g_i > 0$ when $1 - k^2(\lambda_i + k)^{-2}(1-d)^2 < 1$ and this will happen if $\frac{k^2(1-d)^2}{(\lambda_i+k)^2} < 1$ and that is possible if $k \quad (1-d) \quad < (\lambda_i + k)$ . From this formula if $-\infty < d < 0$ then $g_i$ is positive number when $k < \frac{\lambda_i}{-d}$. Now we can state with the following theorem:

***Theorem***:

*Suppose that* $-\infty < d < 0$ , *when* $k < \frac{\lambda_i}{-d}$ *then D will be p.d. The* $\hat{\alpha}_{AUM(k,d)}$ *is superior to the* $\hat{\alpha}_{ME}$ *in the MSE sense* , namely $\Delta$ if and only if $B_1 D^{-1} B_1' < \sigma^2$.

## 4 Conclusion

In this paper, the two parameters estimator under stochastic linear restriction model is proposed. The proposed estimator is coming by two stages: one by introducing a new estimator AU(k-d) to reduce the bias of the (k-d) estimator that is given from Özkale and Kaçiranlar (2007) and second by using the stochastic linear restrictions we got the final proposed estimator AUM(k-d) as an alternative to the mixed estimator ME and then we discussed the properties of the proposed estimator. Finally, sufficient and necessary conditions for the superiority of the AUM(k-d) estimator over the ME in terms of mean squared errors matrix are established.

## References


- Akdeniz, F., Kaciranlar, S. (1995). On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE. *Communications in Statistics-Theory and Methods,* **24**,1789-1797.
- Farebrother, R. W. (1976). Fruther results on the mean square error of ridge regression. *Journal of Royal statistical Society*, **B.38**, 284-250**.**
- Hoerl, A. E. and Kennard, R. W. (1970a). Ridge Regression : Biased estimation for non-orthogonal problem. *Technometrics*, **12**, 55-67.
- Hoerl, A. E. and Kennard, R. W. (1970b). Ridge Regression: Application for non-orthogonal problem. *Technometrics*, **12**, 69-82.
- Kaciranlar S., Akdeniz, S. S. F., Styan, G. P. H.and Werner, H. J., (1999). A new

biased estimator in linear regression and detailed analysis of the widely-analysed dataset on Portland Cement. *Sankhya* **B**, **61**, 443-459.

- Kadiyala, K., (1984). A class almost unbiased and efficient estimators of regression coefficients. *Economics Letter*, **16**, 293-296.
- Liu, K. (1993). A new class of biased estimate in linear regression. *Communications in Statistics-Theory and Methods*, **22**, 393-402.
- Ohtani, K., (1986). On small sample properties of the almost unbiased generalized ridge estimator. *Communications in Statistics-Theory and Methods*, **15**, 1571-1578.
- Ozkale, M. Revan and Kaciranlar, Selahattin (2007). The Restricted and Unrestricted Two-Parameter Estimators. *Communications in Statistics-Theory and Methods*, **36**: 2707–2725.
- Rao, Radhakrishna C., Toutenburg, H., Shalabh and Heumann, C. (2008). *Linear Models and Generalizations Least Squares and Alternatives*, Third Extended Edition, Springer.
- Swindel, B. F. (1976). Good ridge estimators based on prior information. *Communications in Statistics-Theory and Methods*, **A5**,1065-1075.
- Singh,B., Chaubey, Y. P., Dweivedi,T. D. (1986). An almost unbiased ridge estimator. *Sankhya* **B**,48, 342-346.
- Theil, H. (1963). On the use of incomplete prior information in regression analysis. *Journal of the American Statistical Association*, **58**, 401-414.
- Theil, H. and Goldberger, A. S. (1961). On pure and mixed estimation in econometrics. *International Economic Review*, 2, 65-78.