# Comparison of Two Estimators of the Regression Coefficient Vector Under Pitman's Closeness Criterion

Jibo Wu [1]

*Chongqing Key Laboratory of Complex Data Analysis & Artificial Intelligence,
Chongqing University of Arts and Sciences, Chongqing*

**Abstract.** Schaffrin and Toutenburg [1] proposed a weighted mixed estimation based on the sample information and the stochastic prior information, and they also show that the weighted mixed estimator is superior to the ordinary least squares estimator under the mean squared error criterion. However, there has no paper to discuss the performance of the two estimators under the Pitman's closeness criterion. This paper presents the comparison of the weighted mixed estimator and the ordinary least squares estimator using the Pitman's closeness criterion. A simulation study is performed to illustrate the performance of the weighted mixed estimator and the ordinary least squares estimator under the Pitman's closeness criterion.

**Keywords.** Ordinary least squares estimator, Pitman's closeness criterion, Weighted mixed estimator

## 1. Introduction

In this paper, we discuss the following multiple linear regression model

$$Y = X\beta + \varepsilon \tag{1}$$

where $Y$ shows an $n \times 1$ vector of observation, $X$ denotes an $n \times p$ known matrix of rank $p$, $\beta$ shows a $p \times 1$ vector of unknown parameters, $\varepsilon$ defines an $n \times 1$ vector of disturbances with $E(\varepsilon) = 0$ and $Cov(\varepsilon) = \sigma^2 I_n$.

By the Gauss-Markov theorem, the classical ordinary least squares (OLS) estimator is given as follows:

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'Y \tag{2}$$

Besides the model (1), we suppose that we have the following stochastic linear restrictions

$$r = R\beta + e, \ e \sim N(0, \sigma^2 V) \tag{3}$$

---

[1]Corresponding Author: Chongqing Key Laboratory of Complex Data Analysis & Artificial Intelligence, Chongqing University of Arts and Sciences, Chongqing 402160, China. Email:linfen52@126.com

where $R$ shows a $j \times p$ known matrix with $rank(R) = j$, $e$ shows a $j \times 1$ vector of disturbances with $E(e) = 0$ and $Cov(e) = \sigma^2 V$, $V$ is assumed to be a known and positive definite matrix, and we suppose that $E(r) = R\beta$. And we also assumed that the random vector $\varepsilon$ is stochastically independent of $e$ [2].

Using (1) and (3), Durbin [3], Theil and Goldberger [4] and Theil [5] have presented the mixed estimator (ME):

$$\hat{\beta}_{ME} = (X'X + R'V^{-1}R)^{-1}(X'Y + R'V^{-1}r) \tag{4}$$

When the sample information given by (1) and the prior information depicted by (3) are to be assigned not necessarily equal weights on the basis of some extraneous considerations in the estimation of regression parameters, Schaffrin and Toutenburg [1] introduced a weighted mixed estimator (WME)

$$\hat{\beta}(w) = (X'X + wR'V^{-1}R)^{-1}(X'Y + wR'V^{-1}r), \ 0 \le w \le 1 \tag{5}$$

The weighted mixed estimator is unbiased estimator.

If we put $w = 0$ in (5), the estimator can be written as

$$\hat{\beta}(0) = (X'X)^{-1}X'Y = \hat{\beta}_{OLS}$$

which is the ordinary least squares estimator.

If we put $w = 1$ in (5), the estimator reduces to

$$\hat{\beta}(1) = (X'X + R'V^{-1}R)^{-1}(X'Y + R'V^{-1}r) = \hat{\beta}_{ME}$$

which is the mixed estimator proposed by Durbin [3], Theil and Goldberger [4] and Theil [5]. This estimator give equal weight to sample (1) and prior information (3).

Although mean squared error (MSE) has been regards as the best criterion for comparing different estimators, Pitman's [6] closeness (PC) criterion has got a great deal of attention in recent years. This criterion is to calculate the probability that one estimator is closer than another to an unknown parameter of interest. Although PC is a meaningful alternative criterion to MSE, it never got widespread acceptance among statisticians, since it is difficulties to compute it. Since Keating and Mason [7] and Rao [8], Rao et al. [9], PC criterion has got considerable attention as a method for comparing the estimators [10].

Many authors have compared estimators using PC criterion in linear regression model. For example, Wencheko [11] compared some shrinkage estimator using PC criterion. Reif [12] compared general pre-test estimator with some regression estimator under PC criterion. Yang et al. [13] compared united biased estimators in linear model. Özkale and Kaçıranlar (2008) [10,14] and Li et al. [15] compared $r - k$ class estimator with ordinary least squares estimator under PC criterion. Wu [16] compared the estimators in two normal linear models for some of the identical parameters under Pitman's closeness criterion.

We will compare the weighted mixed estimator to the OLS estimator under the PC criterion in this paper. A simulation study is presented to illustrate that the weighted mixed estimator is always superior to the OLS estimator.

## 2. Main results

Firstly we list some definitions.

**Definition 2.1.** Suppose that $\hat{\theta}_1$ and $\hat{\theta}_2$ be two estimators of the unknown p-dimensional vector $\theta$. The PC of $\hat{\theta}_1$ relative to $\hat{\theta}_2$ in estimating $\theta$ under a loss function $L(.,\theta)$ is defined as $PMC(\hat{\theta}_1,\hat{\theta}_2,\theta) = P_r(\hat{\theta}_1,\hat{\theta}_2,\theta) = P_r(\Delta(\hat{\theta}_1,\hat{\theta}_2) \geq 0)$, where

$$\Delta(\hat{\theta}_1,\hat{\theta}_2) = L(\hat{\theta}_2,\theta) - L(\hat{\theta}_1,\theta) \tag{6}$$

In this paper, we conisder the following quadratic loss function $L(\hat{\theta},\theta) = (\hat{\theta} - \theta)'U(\hat{\theta} - \theta)$, where $U$ is a given positive definite matrix.

**Definition 2.2.** $\hat{\theta}_1$ is said to be better than $\hat{\theta}_2$ for all $\theta \in \Theta$ in PC if

$$PMC(\hat{\theta}_1,\hat{\theta}_2,\theta) = P_r(\hat{\theta}_1,\hat{\theta}_2,\theta) = P_r(\Delta(\hat{\theta}_1,\hat{\theta}_2) \geq 0) \geq \frac{1}{2}, \; for \; all \; \theta \in \Theta \tag{7}$$

Then we give the comparison results of the two estimators.

**Theorem 2.1.** The PC of $\hat{\beta}(w)$ relative to $\hat{\beta}_{OLS}$ is given by

$$PMC(\hat{\beta}(w),\hat{\beta}_{OLS},\beta)$$

$$= P_r\{\varepsilon_1'[(I_j + wM_j)^{-2} - I_j]\varepsilon_1 + w\varepsilon_1'(I_j + wM_j)^{-2}M_j^{\frac{1}{2}}e_1$$

$$+ we_1'(I_j + wM_j)^{-2}M_j^{\frac{1}{2}}\varepsilon_1 + w^2 e_1'(I_j + wM_j)^{-2}M_j e_1 \leq 0\} \tag{8}$$

**Proof** As matrices $X'X$ and $R'V^{-1}R$ are positive definite, they can be diagonalizable simultaneously, that is, there exists a reversible matrix $Q$ such that

$$Q'X'XQ = I_p \tag{9}$$

$$Q'R'V^{-1}RQ = diag(m_1,...,m_j,0,...,0) = M_p \tag{10}$$

where $m_1,...,m_j$ denote the positive eigenvalues of matrix $R'V^{-1}R(X'X)^{-1}$.

Denote:

$$Q'X'\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \tag{11}$$

$$M_j = diag(m_1,...,m_j) \tag{12}$$

and

$$(M_p^+)^{\frac{1}{2}}Q'R'V^{-1}e = \begin{pmatrix} e_1 \\ 0 \end{pmatrix} \tag{13}$$

where $\varepsilon_1$ and $e_1$ denote j-dimension column random vector and $M_p^+$ denotes Moore-Penrose inverse of matrix $M_p$. Now let $U = X'X$. Thus, we can get:

$$L(\hat{\beta}_{OLS},\beta) = (\hat{\beta}_{OLS} - \beta)'X'X(\hat{\beta}_{OLS} - \beta) = (Q'X'\varepsilon)'Q'X'\varepsilon \tag{14}$$

and

$$L(\hat{\beta}(w), \beta) = (\hat{\beta}(w) - \beta)' X' X (\hat{\beta}(w) - \beta)$$
$$= [(X'X + wR'V^{-1}R)^{-1}(X'Y + wR'V^{-1}r) - \beta]' X' X$$
$$\times [(X'X + wR'V^{-1}R)^{-1}(X'Y + wR'V^{-1}r) - \beta]$$
$$= (Q'X'\varepsilon + wQ'R'V^{-1}e)'(I_p + wM_p)^{-2}(Q'X'\varepsilon + wQ'R'V^{-1}e) \quad (15)$$

Then we can obtain

$$PMC(\hat{\beta}(w), \hat{\beta}_{OLS}, \beta)$$
$$= P_r(L(\hat{\beta}(w), \beta) \le L(\hat{\beta}_{OLS}, \beta))$$
$$= P_r\{(Q'X'\varepsilon + wQ'R'V^{-1}e)'(I_p + wM_p)^{-2}(Q'X'\varepsilon + wQ'R'V^{-1}e)$$
$$\le (Q'X'\varepsilon)'Q'X'\varepsilon\}$$
$$= P_r\{(Q'X'\varepsilon)'(I_p + wM_p)^{-2}Q'X'\varepsilon + w(Q'X'\varepsilon)'(I_p + wM_p)^{-2}(Q'R'V^{-1}e)$$
$$+ w(Q'R'V^{-1}e)'(I_p + wM_p)^{-2}Q'X'\varepsilon + w^2(Q'R'V^{-1}e)'(I_p + wM_p)^{-2}$$
$$\times (Q'R'V^{-1}e) \le (Q'X'\varepsilon)'Q'X'\varepsilon\} \quad (16)$$

Thus, by (11)-(14), we have

$$(Q'X'\varepsilon)'(I_p + wM_p)^{-2}(Q'R'V^{-1}e) = \varepsilon_1'(I_j + wM_j)^{-2}M_j^{\frac{1}{2}}e_1 \quad (17)$$

$$(Q'R'V^{-1}e)'(I_p + wM_p)^{-2}(Q'R'V^{-1}e) = e_1'(I_j + wM_j)^{-2}M_j e_1 \quad (18)$$

$$(Q'R'V^{-1}e)'(I_p + wM_p)^{-2}(Q'X'\varepsilon) = e_1'(I_j + wM_j)^{-2}M_j^{\frac{1}{2}}\varepsilon_1 \quad (19)$$

$$(Q'X'\varepsilon)'[(I_p + wM_p)^{-2} - I_p](Q'X'\varepsilon) = \varepsilon_1'[(I_j + wM_j)^{-2} - I_j]\varepsilon_1 \quad (20)$$

Thus put (17)-(20) into (16), we have

$$PMC(\hat{\beta}(w), \hat{\beta}_{OLS}, \beta)$$
$$= P_r\{\varepsilon_1'[(I_j + wM_j)^{-2} - I_j]\varepsilon_1 + w\varepsilon_1'(I_j + wM_j)^{-2}M_j^{\frac{1}{2}}e_1$$
$$+ we_1'(I_j + wM_j)^{-2}M_j^{\frac{1}{2}}\varepsilon_1 + w^2 e_1'(I_j + wM_j)^{-2}M_j e_1 \le 0\} \quad (21)$$

**Theorem 2.2.** $\hat{\beta}(w)$ always dominates $\hat{\beta}_{OLS}$ in the Pitman's closeness criterion for $0 < w < 1$.

**Proof** By Theorem 2.1, we have

$$PMC(\hat{\beta}(w), \hat{\beta}_{OLS}, \beta)$$
$$= P_r\{\varepsilon_1'[(I_j + wM_j)^{-2} - I_j]\varepsilon_1 + w\varepsilon_1'(I_j + wM_j)^{-2}M_j^{\frac{1}{2}}e_1$$

$$+we_1'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}\varepsilon_1 + w^2e_1'(I_j+wM_j)^{-2}M_je_1 \leq 0\}$$

$$= P_r\{-\varepsilon_1'(I_j+wM_j)^{-2}(2wM_j+w^2M_j^2)\varepsilon_1 + w\varepsilon_1'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}e_1$$

$$+we_1'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}\varepsilon_1 + w^2e_1'(I_j+wM_j)^{-2}M_je_1 \leq 0\}$$

$$\geq P_r\{-w\varepsilon_1'(I_j+wM_j)^{-2}M_j\varepsilon_1 + \varepsilon_1'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}e_1$$

$$+e_1'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}\varepsilon_1 + we_1'(I_j+wM_j)^{-2}M_je_1 \leq 0\} \tag{22}$$

Now we calculate:

$$P_r\{-w\varepsilon_1'(I_j+wM_j)^{-2}M_j\varepsilon_1 + \varepsilon_1'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}e_1$$

$$+e_1'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}\varepsilon_1 + we_1'(I_j+wM_j)^{-2}M_je_1 \leq 0\} \tag{23}$$

Denote $\eta_1 = e_1$ and $\eta_2 = -\varepsilon_1$. Then (23) becomes :

$$P_r\{-w\eta_2'(I_j+wM_j)^{-2}M_j\eta_2 - \eta_2'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}\eta_1$$

$$-\eta_1'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}\eta_2 + w\eta_1'(I_j+wM_j)^{-2}M_j\eta_1 \leq 0\} \tag{24}$$

Since $\varepsilon_1 \sim N(0,\sigma^2 I_j)$, $e_1 \sim N(0,\sigma^2 I_j)$, and $\varepsilon_1$ is independent of $e_1$. We have: $\eta_1 \sim N(0,\sigma^2 I_j)$, $\eta_2 \sim N(0,\sigma^2 I_j)$, and $\eta_1$ is independent of $\eta_2$, thus we can change $\varepsilon_1$ as $\eta_1$ and change $e_1$ as $\eta_2$ in (23), we obtain

$$P_r\{-w\eta_1'(I_j+wM_j)^{-2}M_j\eta_1 + \eta_1'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}\eta_2$$

$$+\eta_2'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}\eta_1 + w\eta_2'(I_j+wM_j)^{-2}M_j\eta_2 \leq 0\}$$

$$= 1 - P_r\{-w\eta_2'(I_j+wM_j)^{-2}M_j\eta_2 - \eta_2'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}\eta_1$$

$$-\eta_1'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}\eta_2 + w\eta_1'(I_j+wM_j)^{-2}M_j\eta_1 \leq 0\} \tag{25}$$

On the other hand (24) and (25) is equivalent to (23), we have

$$P_r\{-w\eta_1'(I_j+wM_j)^{-2}M_j\eta_1 + \eta_1'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}\eta_2$$

$$+\eta_2'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}\eta_1 + w\eta_2'(I_j+wM_j)^{-2}M_j\eta_2 \leq 0\}$$

$$= P_r\{-w\eta_2'(I_j+wM_j)^{-2}M_j\eta_2 - \eta_2'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}\eta_1$$

$$-\eta_1'(I_j+wM_j)^{-2}M_j^{\frac{1}{2}}\eta_2 + w\eta_1'(I_j+wM_j)^{-2}M_j\eta_1 \leq 0\} = \frac{1}{2} \tag{26}$$

Thus:

$$PMC(\hat{\beta}(w),\hat{\beta}_{OLS},\beta) \geq \frac{1}{2} \tag{27}$$

**Table 1.** The PMC of WME relative to OLS when $n = 50$

| $\gamma = 0.9$ | $w = 0.1$ | $w = 0.2$ | $w = 0.3$ | $w = 0.4$ | $w = 0.5$ |
|---|---|---|---|---|---|
| | 0.8084 | 0.7998 | 0.7979 | 0.7896 | 0.7907 |
| | $w = 0.6$ | $w = 0.7$ | $w = 0.8$ | $w = 0.9$ | $w = 1.0$ |
| | 0.7793 | 0.7672 | 0.7769 | 0.7687 | 0.7680 |
| $\gamma = 0.99$ | $w = 0.1$ | $w = 0.2$ | $w = 0.3$ | $w = 0.4$ | $w = 0.5$ |
| | 0.8690 | 0.8530 | 0.8468 | 0.8336 | 0.8279 |
| | $w = 0.6$ | $w = 0.7$ | $w = 0.8$ | $w = 0.9$ | $w = 1.0$ |
| | 0.8262 | 0.8107 | 0.8130 | 0.8042 | 0.8044 |
| $\gamma = 0.999$ | $w = 0.1$ | $w = 0.2$ | $w = 0.3$ | $w = 0.4$ | $w = 0.5$ |
| | 0.8960 | 0.8849 | 0.8866 | 0.8794 | 0.8805 |
| | $w = 0.6$ | $w = 0.7$ | $w = 0.8$ | $w = 0.9$ | $w = 1.0$ |
| | 0.8800 | 0.8771 | 0.8789 | 0.8779 | 0.8728 |

The proof of Theorem 2.2 is completed.

**Remark 2.1.** From Theorem 2.2, we can see that the weighted mixed estimator is superior over the OLS estimator in the PC criterion. In fact, these results are easy to be understand. Because by adding the prior information, we have got more information about the unknown regression parameters, and naturally hope that the new estimator may have good accuracy.

## 3. A simulation study

In this section, we do some simulation to illustrate the performance of the weighted mixed estimator (WME) over the OLS estimator under the PC criterion.

By the methods proposed by Liu [17] and Kibria [18], we have

$$x_{ij} = (1 - \gamma^2)^{1/2} w_{ij} + \gamma w_{ip}, \ i = 1, 2, ..., n, \ j = 1, 2, ..., 4 \tag{28}$$

and

$$y_i = (1 - \gamma^2)^{1/2} w_i + \gamma w_{i4}, \ i = 1, 2, ..., n, \ j = 1, 2, ..., 4 \tag{29}$$

where $w_{ij}$ are independent standard normal pseudo-random numbers and the correlation between any two explanatory variables is given by $\gamma^2$. We consider $\gamma = 0.9, \ 0.99$ and $0.999$, and we choose $n = 50$ and $100$, $\sigma^2 = 1$.

Let us discuss the following restrictions

$$r = R\beta + e, R = (1 - 1 - 1 \ 0), \ r = 1, \ e \sim N(0, \hat{\sigma}^2) \tag{30}$$

For Tables 1–2, we can see that the PC of WME relative to OLS estimator is greater than 0.5, that is to say the WME is better than the OLS estimator in the PC criterion.

**Table 2.** The PMC of WME relative to OLS when $n = 100$

| $\gamma = 0.9$ | $w = 0.1$ | $w = 0.2$ | $w = 0.3$ | $w = 0.4$ | $w = 0.5$ |
|---|---|---|---|---|---|
| | 0.8626 | 0.8670 | 0.8587 | 0.8573 | 0.8547 |
| | $w = 0.6$ | $w = 0.7$ | $w = 0.8$ | $w = 0.9$ | $w = 1.0$ |
| | 0.8528 | 0.8528 | 0.8555 | 0.8479 | 0.8469 |
| $\gamma = 0.99$ | $w = 0.1$ | $w = 0.2$ | $w = 0.3$ | $w = 0.4$ | $w = 0.5$ |
| | 0.8362 | 0.8374 | 0.8282 | 0.8196 | 0.8118 |
| | $w = 0.6$ | $w = 0.7$ | $w = 0.8$ | $w = 0.9$ | $w = 1.0$ |
| | 0.8095 | 0.8030 | 0.7952 | 0.7937 | 0.7926 |
| $\gamma = 0.999$ | $w = 0.1$ | $w = 0.2$ | $w = 0.3$ | $w = 0.4$ | $w = 0.5$ |
| | 0.9132 | 0.8922 | 0.8899 | 0.8911 | 0.8850 |
| | $w = 0.6$ | $w = 0.7$ | $w = 0.8$ | $w = 0.9$ | $w = 1.0$ |
| | 0.8858 | 0.8907 | 0.8872 | 0.8810 | 0.8840 |

## 4. Conclusion

In this paper, we make the comparison between the weighted mixed estimator and the ordinary least squares estimator in the PC criterion and we show that the weighted mixed estimator is always superior to the ordinary least squares estimator in the PC criterion. A simulation simulation is given to show that the weighted mixed estimator is better than the ordinary least squares estimator in the PC criterion.

## References

[1] B. Schaffrin and H. Toutenburg, *Weighted mixed regression*, Zeitschrift fur Angewandte Mathematik und Mechanik, **70**(1990), 735–738.

[2] C.R. Rao and H. Toutenburg, *Linear models: least suqres and alternatives*, Springer, New York, 1995.

[3] J. Durbin, *A note on regression when there is extraneous information about one of the coefficients*, J Am Stat Assoc. **48** (1985), 799–808.

[4] H. Theil and A.S. Goldberger, *On pure and mixed statistical extimation in economics*, Int Econ Rev. **2**(1961), 65–78.

[5] H. Theil, On the use of incomplete prior information in regression analysis, J Am Stat Assoc. 58 (1963), 401–414.

[6] E.J.G. Pitman, *The closest estimates of statistical parameters*, Proc Cambridge Philosop Soc. **33**(1937), 212–222.

[7] J. P. Keating and R.L. Mason, *Practical relevance of an alternative criterion in estimation*, Am Stat **39**(1985), 203–205.

[8] C.R. Rao, *Some comments on the minimum mean square error as a criterion of estimation*, Statistics and related topics. North Holland Amsterdam. 1981, 123–143.

[9] C. R. Rao, J.P. Keating and R.L. Mason, The Pitman nearness criterion and determination, Commun Statist Theor Meth. 15(1986), 3173–3191.

[10]  M.R. Özkale and S. Kaçıranlar, Comparisons of the $r - k$ class estimator to the ordinary least squares estimator under the Pitman's closeness criterion, Stat pap 49(2008), 503–512.

[11]  E. Wencheko, *Comparison of regression estimators using Pitman measures of nearness*, Stat. Pap. **42**(2001), 375–386.

[12]  J. Reif, *Pitman closeness in classes of general pre-test and regression estimators*, Commun Statist Theor Meth. **35**(2006), 263–280.

[13]  H. Yang, W.X. Li and J.W. Xu, *Comparison of Two Estimators of Parameters Under Pitman Nearness Criterion*, Commun Statist Theor Meth. **17**(2010), 3081–3094.

[14]  M.R. Özkale and S. Kaçıranlar, *Erratum to: Comparisons of the $r - k$ class estimator to the ordinary least squares estimator under the Pitman's closeness criterion*, Stat pap. **53**(2012), 505–505.

[15]  W.X. Li, H. Yang and J.B. Wu, *Some comments on: M. R. Özkale, S. Kaçıranlar: Comparisons of the $r - k$ class estimator to the ordinary least squares estimator under the Pitman's closeness criterion*, Stat pap. **53**(2012), 497–503.

[16]  J.B. Wu, *Estimation of the parameters in two linear models with some of the identical parameter vectors under the Pitman's closeness criterion*. Journal of the National Science Foundation of Sri Lanka, **47**(2019), 363–368.

[17]  K.J. Liu, *Using Liu-type estimator to combat collinearity*, Commun Stat Theory Methods. **32**(2003), 1009–1020.

[18]  B.M.G. Kibria, *Performance of some new ridge regression estimators*, Comm Stat Simul Comput. **32**(2003), 2389–2413.