Journal of Statistical and Econometric Methods, vol.2, no.3, 2013, 153-174
ISSN: 1792-6602 (print), 1792-6939 (online)
Scienpress Ltd, 2013

The Superiorities of Minimum Bayes Risk Linear Unbiased Estimator in Two Seemingly Unrelated Regressions[1]

Radhey S. Singh[2], Lichun Wang[3] and Huiming Song[4]

Abstract

In the system of two seemingly unrelated regressions, the minimum Bayes risk linear unbiased (MBRLU) estimators of regression parameters are derived. The superiorities of the MBRLU estimators over the classical estimators are investigated, respectively, in terms of the mean square error matrix (MSEM) criterion, the predictive Pitman closeness (PRPC) criterion and the posterior Pitman closeness (PPC) criterion.

Mathematics Subject Classification: 62C10; 62J05
Keywords: Seemingly unrelated regressions; generalized least square estimator; MBRLU estimator; MSEM criterion; PRPC criterion; PPC criterion

---

[1]Sponsored by NSERC of Canada Grant (400045) and Scientific Foundation of BJTU(2012JBM105)
[2] Department of Statistics and Actuarial Sciences, University of Waterloo, Canada. Corresponding author.
[3] Department of Mathematics, Beijing Jiaotong University, Beijing 100044, China.
[4] Department of Statistics and Finance, USTC, Hefei 230026, China.

Article Info: *Received* : March 21, 2013. *Revised* : May 2, 2013
Published online : September 1, 2013

1 Introduction

Consider a system consisting of two seemingly unrelated regressions (SUR)

$$\begin{cases} Y_1 = X_1\beta_1 + \epsilon_1 \\ Y_2 = X_2\beta_2 + \epsilon_2 \end{cases}, \tag{1.1}$$

where $Y_i(i = 1, 2)$ are $n \times 1$ vectors of observations, $X_i(i = 1, 2)$ are $n \times p_i$ matrices with full column rank, $\beta_i(i = 1, 2)$ are $p_i \times 1$ vectors of unknown regression parameters, $\epsilon_i(i = 1, 2)$ are $n \times 1$ vectors of error variables, and

$$E(\epsilon_i) = 0, \quad Cov(\epsilon_i, \epsilon_j) = \sigma_{ij} I_n, \quad i, j = 1, 2,$$

where $\Sigma^* = (\sigma_{ij})$ is a 2×2 non-diagonal positive definite matrix. This kind of system has been widely applied in many fields such as econometrics, social and biological sciences and so on. It was first introduced to statistics by the Zellner's pioneer works (Zellner (1962, 1963)) and later developed by Kementa and Gilbert (1968), Revankar (1974), Mehta and Swamy (1976), Schmidt (1977), Wang (1989) and Lin (1991), etc.

Denote $Y = (y_1', y_2')'$, $X = \text{diag}(X_1, X_2)$, $\beta = (\beta_1', \beta_2')'$, $\epsilon = (\epsilon_1', \epsilon_2')'$. One can represent the system (1.1) as the following regression model

$$Y = X\beta + \epsilon, \quad \epsilon \sim (0, \Sigma^* \otimes I_n), \tag{1.2}$$

where $\otimes$ denotes the Kronecker product operator.

Following from Wang et al (2011), we know that if Σ^* is known then the generalized least square estimator of $\beta_i(i = 1, 2)$ would be

$$\bar{\beta}_{1,GLS} = (X_1'X_1)^{-1}X_1' \left[I_n - \rho^2 \sum_{i=0}^{\infty} (\rho^2 P_2 P_1)^i P_2 N_1 \right] (Y_1 - \frac{\sigma_{12}}{\sigma_{22}} N_2 Y_2) \tag{1.3}$$

and

$$\bar{\beta}_{2,GLS} = (X_2'X_2)^{-1}X_2' \left[I_n - \rho^2 \sum_{i=0}^{\infty} (\rho^2 P_1 P_2)^i P_1 N_2 \right] (Y_2 - \frac{\sigma_{21}}{\sigma_{11}} N_1 Y_1), \tag{1.4}$$

where $P_i = X_i(X_i'X_i)^{-1}X_i'$, $N_i = I_n - P_i$ and $\rho^2 = \sigma_{12}^2/(\sigma_{11}\sigma_{22})$.

Further, also by the Theorem 3.1 of Wang et al (2011), we know that under the condition that $X_1'X_2 = 0$ (see Zellner (1963)) or $X_1 = (X_2, L)$ (see Revankar (1974)) or $P_1P_2P_1N_2 = 0$ (see Liu (2002)), formally $\bar{\beta}_{1,GLS}$ and $\bar{\beta}_{2,GLS}$ have unique simpler form, respectively. Thus, in what follows we assume $X_1'X_2 = 0$ which implies $X_i'N_j = 0 (i \neq j)$ and causes $\bar{\beta}_{1,GLS}$ and $\bar{\beta}_{2,GLS}$ to be simplified into

$$\hat{\beta}_1^* = \hat{\beta}_1 - \frac{\sigma_{12}}{\sigma_{22}}(X_1'X_1)^{-1}X_1'Y_2, \quad (1.5)$$

$$\hat{\beta}_2^* = \hat{\beta}_2 - \frac{\sigma_{12}}{\sigma_{11}}(X_2'X_2)^{-1}X_2'Y_1, \quad (1.6)$$

where

$$\hat{\beta}_1 = (X_1'X_1)^{-1}X_1'Y_1, \quad \hat{\beta}_2 = (X_2'X_2)^{-1}X_2'Y_2.$$

In Section 2 we derive the Bayes minimum risk linear unbiased (MBRLU) estimator for β and accordingly the MBRLU estimators for $\beta_i (i = 1, 2)$. In Section 3 the superiorities of the MBRLU estimators of $\beta_i (i = 1, 2)$ are established based on the mean square error matrix (MSEM) criterion. In Section 4 we discuss the superiorities of MBRLU estimators in terms of the predictive Pitman closeness (PRPC) criterion and the posterior Pitman closeness (PPC) criterion, respectively. In the case that the design matrices are non-full rank, we investigate the superiorities of BMRLU estimators of some estimable functions. Finally, brief concluding remarks are made in Section 6.

2 The BMRLU Estimators of Regression Parameters

Normally, there are two different approaches concerned with Bayes estimation in linear model. The first one supposes that the prior of regression parameter is normal which implies under the normal linear model the posterior is still a normal distribution. Thus under the quadratic loss the Bayes estimator of the regression parameter would be the posterior mean (see Box and Tiao (1973), Berger (1985) and Wang and Chow (1994), etc). Recently, Wang and Veraverbeke (2008) employed this approach to exhibit the superiorities of Bayes and empirical Bayes estimators in two generalized SURs. The second approach proposed by Rao (1973) which yields the minimum Bayes risk linear (MBRL) estimator of the regression parameter by minimizing the Bayes risk under the assumption that some moment conditions of the prior are given. Rao (1976) further pointed out that the admissible linear estimators of regression parameter are either MBRL estimators or the limit of MBRL estimators. Gruber (1990) proposed a MBRL estimator for an estimable function of the regression parameter and obtained an alternative form of MBRL estimator. Some results related to this area can be found in Trenkler and Wei (1996), Zhang (2005) and others. In this paper, we use the second approach to derive MBRL estimator of the regression parameter and discuss the superiorities of MBRL estimator in terms of MSEM, PRPC and PPC criterion, respectively.

Denote the prior of β by $\pi(\beta)$. It is assumed that the prior $\pi(\beta)$ satisfies:

$$E(\beta) = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \hat{=} \mu, \quad Cov(\beta) = \begin{pmatrix} \tau_1^2 I_{p_1} & 0 \\ 0 & \tau_2^2 I_{p_2} \end{pmatrix} \hat{=} V, \tag{2.1}$$

where μ_i and $\tau_i (i = 1, 2)$ are known.

Let the loss function be defined by

$$L(\delta, \beta) = (\delta - \beta)'(\delta - \beta), \tag{2.2}$$

and the linear estimator class of β be

$$\mathcal{F} = \left\{ \widetilde{\beta} = AY + b : \text{where } A \text{ is } (p_1 + p_2) \times 2n \text{ matrix}, \ b \text{ is } (p_1 + p_2) \times 1 \text{ vector} \right\}. \tag{2.3}$$

Then the MBRLU estimator $\hat{\beta}_B$ is defined to minimize the Bayes risk

$$R(\hat{\beta}_B, \beta) = \min_{A,b} R(\widetilde{\beta}, \beta) = \min_{A,b} E[(\widetilde{\beta} - \beta)'(\widetilde{\beta} - \beta)] \tag{2.4}$$

and subject to the constraint $E(\widetilde{\beta} - \beta) = 0$, where E denotes the expectation with respect to (w.r.t.) the joint distribution of (Y, β).

From the constraint, we have $b = (I - AX)\mu$. Note that the fact that

$$\begin{aligned} R(\widetilde{\beta}, \beta) &= E\left\{[AY + (I - AX)\mu - \beta]^{'}[AY + (I - AX)\mu - \beta]\right\} \\ &= E\left\{[A(Y - X\mu) - (\beta - \mu)]^{'}[A(Y - X\mu) - (\beta - \mu)]\right\} \\ &= Etr\left\{[A(Y - X\mu) - (\beta - \mu)][A(Y - X\mu) - (\beta - \mu)]^{'}\right\} \\ &= tr\left\{A(XVX^{'} + \Phi)A^{'} + V - AXV - VX^{'}A^{'}\right\}, \end{aligned}$$

by solving $\frac{\partial R(\widetilde{\beta},\beta)}{\partial A} = 0$, we obtain

$$A = VX'(XVX' + \Phi)^{-1}. \tag{2.5}$$

By the fact that

$$(P + BCB')^{-1} = P^{-1} - P^{-1}B(C^{-1} + B'P^{-1}B)^{-1}B'P^{-1}, \tag{2.6}$$

we obtain

$$A = VX'(XVX' + \Phi)^{-1} = (X'\Phi^{-1}X + V^{-1})^{-1}X'\Phi^{-1}, \tag{2.7}$$

and

$$I - AX = I - (X'\Phi^{-1}X + V^{-1})^{-1}X'\Phi^{-1}X = (X'\Phi^{-1}X + V^{-1})^{-1}V^{-1}. \tag{2.8}$$

Hence, we have

$$\begin{aligned} \hat{\beta}_B &= AY + b = AY + (I - AX)\mu \\ &= (X'\Phi^{-1}X + V^{-1})^{-1}(X'\Phi^{-1}X\hat{\beta}_{LS} + V^{-1}\mu) \\ &= \hat{\beta}_{LS} - (X'\Phi^{-1}X + V^{-1})^{-1}V^{-1}(\hat{\beta}_{LS} - \mu) \\ &= \begin{pmatrix} \hat{\beta}_1^* - (\tau_1^2\sigma_{11.2}^{-1}X_1'X_1 + I_{p_1})^{-1}(\hat{\beta}_1^* - \mu_1) \\ \hat{\beta}_2^* - (\tau_2^2\sigma_{22.1}^{-1}X_2'X_2 + I_{p_2})^{-1}(\hat{\beta}_2^* - \mu_2) \end{pmatrix}. \end{aligned} \tag{2.9}$$

Thus the MBRLU estimators of β_i $(i = 1, 2)$ are

$$\hat{\beta}_{1B} = \hat{\beta}_1^* - (\tau_1^2 \sigma_{11.2}^{-1} X_1^{'} X_1 + I_{p_1})^{-1} (\hat{\beta}_1^* - \mu_1), \tag{2.10}$$

$$\hat{\beta}_{2B} = \hat{\beta}_2^* - (\tau_2^2 \sigma_{22.1}^{-1} X_2^{'} X_2 + I_{p_2})^{-1} (\hat{\beta}_2^* - \mu_2), \tag{2.11}$$

where $\sigma_{11.2} = \sigma_{11} - \sigma_{12}^2 \sigma_{22}^{-1}$ and $\sigma_{22.1} = \sigma_{22} - \sigma_{12}^2 \sigma_{11}^{-1}$.

3 The Superiorities of MBRLU Estimator Under MSEM Criterion

We state the following MESM superiorities of $\hat{\beta}_{iB} (i = 1, 2)$.

Theorem 3.1 Let the GLS estimators and MBRLU estimators of β_i are given by (1.5),(1.6) and (2.10), (2.11) respectively, then

$$M(\hat{\beta}_i^*) - M(\hat{\beta}_{iB}) > 0, \quad i = 1, 2.$$

Proof: We only prove the above conclusion for the case of $i = 1$. Denote $B_1 = (\tau_1^2 \sigma_{11.2}^{-1} X_1^{'} X_1 + I_{p_1})^{-1}$, we have

$$\begin{aligned} M(\hat{\beta}_{1B}) &= E\left[(\hat{\beta}_{1B} - \beta_1)(\hat{\beta}_{1B} - \beta_1)^{'}\right] \\ &= E\left\{\left[(\hat{\beta}_1^* - \beta_1) - B_1(\hat{\beta}_1^* - \mu_1)\right]\left[(\hat{\beta}_1^* - \beta_1) - B_1(\hat{\beta}_1^* - \mu_1)\right]^{'}\right\} \\ &= M(\hat{\beta}_1^*) - E[(\hat{\beta}_1^* - \beta_1)(\hat{\beta}_1^* - \mu_1)^{'}]B_1^{'} - B_1 E[(\hat{\beta}_1^* - \mu_1)(\hat{\beta}_1^* - \beta_1)^{'}] \\ &\quad + B_1 E\left[(\hat{\beta}_1^* - \mu_1)(\hat{\beta}_1^* - \mu_1)^{'}\right] B_1^{'} \\ &= M(\hat{\beta}_1^*) - J_1 B_1^{'} - B_1 J_1^{'} - B_1 J_2 B_1^{'}, \end{aligned} \tag{3.1}$$

where

$$\begin{aligned} J_2 &= Cov(\hat{\beta}_1^*) = E\left\{Cov\left(\hat{\beta}_1^* | \beta\right)\right\} + cov\left\{E\left(\hat{\beta}_1^* | \beta\right)\right\} \\ &= E\left\{cov\left(\hat{\beta}_1^* \mid \beta\right)\right\} + \tau_1^2 I, \end{aligned} \tag{3.2}$$

and

$$\begin{aligned} E\left\{Cov\left(\hat{\beta}_1^*|\beta_1\right)\right\} &= E\left\{E\left[\left(\hat{\beta}_1^*-\beta_1\right)\left(\hat{\beta}_1^*-\beta_1\right)'|\beta_1\right]\right\} \\ &= E\left\{Cov(\hat{\beta}_1|\beta_1)+Cov\left(\sigma_{12}\sigma_{22}^{-1}(X_1'X_1)^{-1}X_1'Y_2\ |\beta_1\right)\right. \\ &\quad \left.-2Cov\left(\hat{\beta}_1,\sigma_{12}\sigma_{22}^{-1}(X_1'X_1)^{-1}X_1'Y_2\ |\beta_1\right)\right\} \\ &= \sigma_{11}(X_1'X_1)^{-1}+\sigma_{12}^2\sigma_{22}^{-1}(X_1'X_1)^{-1}-2\sigma_{12}^2\sigma_{22}^{-1}(X_1'X_1)^{-1} \\ &= \left(\sigma_{11}-\sigma_{12}^2\sigma_{22}^{-1}\right)(X_1'X_1)^{-1}=\sigma_{11.2}(X_1'X_1)^{-1}. \qquad (3.3) \end{aligned}$$

Putting (3.3) into (3.2) we have

$$J_2=E\left\{Cov\left(\hat{\beta}_1^*|\beta_1\right)\right\}+\tau_1^2I=\sigma_{11.2}(X_1'X_1)^{-1}+\tau_1^2I. \qquad (3.4)$$

Then J_1 can be expressed as follows

$$J_1=E\left[(\hat{\beta}_1^*-\beta_1)(\hat{\beta}_1^*-\mu_1)'\right]=J_2-Cov(\beta_1)=\sigma_{11.2}(X_1'X_1)^{-1}. \qquad (3.5)$$

Combining (3.4) and (3.5) with (3.1), we obtain

$$\begin{aligned} M(\hat{\beta}_{1B})-M(\hat{\beta}_1^*) &= J_1B_1'+B_1J_1'-B_1J_2B_1' \\ &= B_1\left[B_1^{-1}\sigma_{11.2}(X_1'X_1)^{-1}+\sigma_{11.2}(X_1'X_1)^{-1}B_1^{-1}\right. \\ &\quad \left.-\tau_1^2I-\sigma_{11.2}(X_1'X_1)^{-1}\right]B_1' \\ &= B_1\left[\tau_1^2I+\sigma_{11.2}(X_1'X_1)^{-1}\right]B_1'>0. \qquad (3.6) \end{aligned}$$

Theorem 3.1 has been proved.

Obviously, the MSEM is much stronger than the MSE. A point estimator could be MSE superior to another, but not necessarily superior in sense of MSEM. Hence, we have

$$MSE(\hat{\beta}_i^*)-MSE(\hat{\beta}_{iB})>0, \quad i=1,2.$$

4 Superiorities of MBRLU Estimators Under PRPC and PPC Criterion

The criterion of Pitman Closeness (PC), originally introduced by Pitman (1937), is based on the probabilities of the relative closeness of competing estimators to an unknown parameter or parameter vector. After a long fallow period, renewed interest in this topic has been sparked in the last twenty years. Rao (1981), Keating and Mason (1985) and Rao et al. (1986) helped to resurrect the criterion as an alternative comparison criterion to traditional criterion such as MSE criterion and mean absolute error (MAE) criterion. Mason et al. (1990) and Fountain and Keating (1994) proposed some general methods for the comparisons between linear estimators under PC criterion. Many important contributions to this direction were described by Keating et al. (1993) and others.

Definition 4.1 Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two different estimators of θ, $L(\hat{\theta}, \theta)$ be the loss function. if

$$P(L(\hat{\theta}_1, \theta) \leq L(\hat{\theta}_2, \theta)) \geq 0.5, \ \forall \ \theta \in \Theta,$$

with strict inequality " $>$ " for some $\theta \in \Theta$, the parameter space, then $\hat{\theta}_1$ is said to be Pitman closer than $\hat{\theta}_2$, or $\hat{\theta}_1$ is said to be superior to $\hat{\theta}_2$ under PC criterion.

Ghosh and Sen (1991) introduced two alternative notions of PC motivated from Bayesian viewpoint, which are called Predictive Pitman Closeness (PRPC) and Posterior Pitman Closeness (PPC) criterion. They are defined as follows:

Definition 4.2 Let Γ be a class of prior distributions of θ, $\hat{\theta}_1$ and $\hat{\theta}_2$ be two different estimators of θ, then $\hat{\theta}_1$ is said to be superior to $\hat{\theta}_2$ under PRPC criterion if

$$P_\pi(L(\hat{\theta}_1, \theta) \leq L(\hat{\theta}_2, \theta)) \geq 0.5, \quad \forall \ \pi \in \Gamma,$$

where P_π is computed under the joint distribution of Y and θ for every $\pi \in \Gamma$.

Definition 4.3 Suppose π is a prior distribution of θ, $\hat{\theta}_1$ and $\hat{\theta}_2$ are two different estimators of θ, then $\hat{\theta}_1$ is said to superior to $\hat{\theta}_2$ under PPC criterion if

$$P_\pi(L(\hat{\theta}_1, \theta) \leq L(\hat{\theta}_2, \theta)|y) \geq 0.5, \quad \forall\, y \in \mathcal{Y},$$

with strict inequality " $>$ " for some $y \in \mathcal{Y}$, where $\mathcal{Y}$ is the sample space.

Obviously, if an estimator $\hat{\theta}_1$ of θ is superior to $\hat{\theta}_2$ for every $\pi \in \Gamma$ under PPC criterion, then it is also superior to $\hat{\theta}_2$ under PRPC criterion. The converse is not necessarily true. Ghosh and Sen presented an example to show that the classical James-Stein estimator is superior to the sample mean under PRPC criterion for all priors, but it is not hold under the PPC criterion.

Let the loss function be defined by (2.2) and in this section

$$\varepsilon|\beta \sim N(0, \Sigma^* \otimes I_n). \tag{4.1}$$

For the MBRLU estimator $\hat{\beta}_{1B}$, we have the following results.

Theorem 4.1 Let the GLS estimator and MBRLU estimator of β_1 be given by (1.5) and (2.10). If

$$\frac{\tau_1^2}{\sigma_{11.2}} \leq \frac{\lambda_{p_1}(p_1 - 2)}{2p_1\lambda_1^2}, \tag{4.2}$$

then

$$P_\pi\left(L(\hat{\beta}_{1B}, \beta_1)) \leq L(\hat{\beta}_1^*, \beta_1)\right) \geq 0.5, \quad \text{for every } \pi \in \Gamma(\beta_1),$$

where λ_1 and λ_{p_1} are the maximum and the minimum eigenvalues of $X_1^{'}X_1$ and $\Gamma(\beta_1) = \{\pi(\beta_1) : E(\beta_1) = \mu_1,\ Cov(\beta_1) = \tau_1^2 I_{p_1}\}$.

Proof: Denote $B_1 = (\tau_1^2\sigma_{11.2}^{-1}X_1^{'}X_1 + I_{p_1})^{-1}$ and $W\left(\hat{\beta}_{1B}, \beta_1^*; \beta_1\right) = L(\hat{\beta}_{1B}, \beta_1) - L(\hat{\beta}_1^*, \beta_1)$.

By (2.10) we have

$$\begin{aligned} L(\hat{\beta}_{1B}, \beta_1) &= [(\hat{\beta}_1^* - \beta_1) - B_1(\hat{\beta}_1^* - \mu_1)]^{'}[(\hat{\beta}_1^* - \beta_1) - B_1(\hat{\beta}_1^* - \mu_1)] \\ &= L(\hat{\beta}_1^*, \beta_1) - 2(\hat{\beta}_1^* - \mu_1)^{'}B_1^{'}(\hat{\beta}_1^* - \beta_1) + (\hat{\beta}_1^* - \mu_1)^{'}B_1^2(\hat{\beta}_1^* - \mu_1). \end{aligned} \tag{4.3}$$

Hence $W < 0$ is equivalent to

$$(\hat{\beta}_1^* - \mu_1)' B_1^2 (\hat{\beta}_1^* - \mu_1) - 2(\hat{\beta}_1^* - \mu_1)' B_1' (\hat{\beta}_1^* - \beta_1) \le 0. \tag{4.4}$$

Since $B_1^2 \le B_1$, we know that (4.4) is implied by

$$(\hat{\beta}_1^* - \mu_1)' B_1 (\hat{\beta}_1^* - \mu_1) - 2(\hat{\beta}_1^* - \mu_1)' B_1 (\hat{\beta}_1^* - \beta_1) \le 0. \tag{4.5}$$

Substituting $\hat{\beta}_1^* - \mu_1 = \hat{\beta}_1^* - \beta_1 - (\mu_1 - \beta_1)$ into (4.5), it is equivalent to

$$(\beta_1 - \mu_1)' B_1 (\beta_1 - \mu_1) \le (\hat{\beta}_1^* - \beta_1)' B_1 (\hat{\beta}_1^* - \beta_1). \tag{4.6}$$

Since

$$\left(\frac{\tau_1^2}{\sigma_{11.2}}\lambda_1 + 1\right)^{-1} I_{p_1} \le B_1 = \left(\frac{\tau_1^2}{\sigma_{11.2}} X_1' X_1 + I_{p_1}\right)^{-1} \le \left(\frac{\tau_1^2}{\sigma_{11.2}}\lambda_{p_1} + 1\right)^{-1} I_{p_1},$$

It is easy to see that inequality (4.6) is implied by

$$\left(\frac{\tau_1^2}{\sigma_{11.2}}\lambda_{p_1} + 1\right)^{-1} (\beta_1 - \mu_1)'(\beta_1 - \mu_1) \le \left(\frac{\tau_1^2}{\sigma_{11.2}}\lambda_1 + 1\right)^{-1} (\hat{\beta}_1^* - \beta_1)'(\hat{\beta}_1^* - \beta_1). \tag{4.7}$$

Since $0 < \frac{\tau_1^2 \sigma_{11.2}^{-1} \lambda_1 + 1}{\tau_1^2 \sigma_{11.2}^{-1} \lambda_{p_1} + 1} \le \frac{\lambda_1}{\lambda_{p_1}}$, So (4.7) is implied by

$$\frac{\lambda_1}{\lambda_{p_1}} (\beta_1 - \mu_1)'(\beta_1 - \mu_1) \le (\hat{\beta}_1^* - \beta_1)'(\hat{\beta}_1^* - \beta_1). \tag{4.8}$$

Note that $(\hat{\beta}_1^* - \beta_1)|\beta \sim N\left(0, \sigma_{11.2}(X_1'X_1)^{-1}\right)$. Let $Z = \sigma_{11.2}^{-1/2}(X_1'X_1)^{1/2}(\hat{\beta}_1^* - \beta_1)$, then $Z|\beta \sim N\left(0, I_{p_1}\right)$. So the inequality(4.8) is equivalent to

$$\frac{\lambda_1}{\lambda_{p_1}} (\beta_1 - \mu_1)'(\beta_1 - \mu_1) \le \sigma_{11.2} Z' (X_1'X_1)^{-1} Z. \tag{4.9}$$

Notice that $(X_1'X_1)^{-1} > \lambda_1^{-1} I_{p_1}$, hence (4.9) is implied by

$$\frac{\lambda_1^2}{\lambda_{p_1}\sigma_{11.2}} ||\beta_1 - \mu_1||^2 \le Z'Z. \tag{4.10}$$

Since $Z'Z|\beta \sim \chi^2_{p_1}$, by (4.4)–(4.10) and Markov inequality we have

$$\begin{aligned}
&P_\pi\left(W\left(\hat{\beta}_{1B}, \beta_1^*; \beta_1\right) \le 0\right) \ge P_\pi\left(Z'Z \ge \frac{\lambda_1^2}{\lambda_{p_1}\sigma_{11.2}}\,||\beta_1-\mu_1||^2\right)\\
&= 1 - P_\pi\left(||\beta_1-\mu_1||^2 \ge \frac{\lambda_{p_1}\sigma_{11.2}}{\lambda_1^2}Z'Z\right)\\
&= 1 - E\left[P_\pi\left(||\beta_1-\mu_1||^2 \ge \frac{\lambda_{p_1}\sigma_{11.2}}{\lambda_1^2}Z'Z\,\Big|\,Z'Z\right)\right]\\
&\ge 1 - E\left(\frac{\lambda_1^2 E\,||\beta_1-\mu_1||^2}{\lambda_{p_1}\sigma_{11.2}Z'Z}\right) = 1 - \frac{\lambda_1^2 tr[Cov(\beta_1)]}{\sigma_{11.2}\lambda_{p_1}}E\left(\frac{1}{Z'Z}\right)\\
&= 1 - \frac{\lambda_1^2 p_1 \tau_1^2}{\sigma_{11.2}\lambda_{p_1}(p_1-2)} \ge \frac{1}{2}.
\end{aligned}$$

The proof of Theorem 4.1 is finished.

For the estimator $\hat{\beta}_{2B}$, we have similar result as below.

Theorem 4.2 Let the GLS estimator and MBRLU estimator of β_2 are given by (1.6) and (2.11). If

$$\frac{\tau_2^2}{\sigma_{22.1}} \le \frac{\lambda_{p_2}^*(p_2-2)}{2p_2\lambda_1^{*2}}, \tag{4.11}$$

then

$$P_\pi\left(L(\hat{\beta}_{2B}, \beta_2)) \le L(\hat{\beta}_2^*, \beta_2)\right) \ge 0.5, \quad \text{for every } \pi \in \Gamma(\beta_2),$$

where $\lambda_1^*, \lambda_{p_2}^*$ are the maximum and the minimum eigenvalue of $X_2'X_2$ and $\Gamma(\beta_2) = \{\pi(\beta_2) : E(\beta_2) = \mu_2,\ Cov(\beta_2) = \tau_2^2 I_{p_2}\}$.

To discuss PPC properties of MBRLU estimators, we further assume the prior $\pi(\beta)$ is the normal distribution,i.e.

$$\beta \sim N(\mu, V). \tag{4.12}$$

Thus we have the following results.

Theorem 4.3 Under the assumptions (4.1) and (4.12).

$$P_\pi\left(L(\hat{\beta}_{iB}, \beta_i)) \le L(\hat{\beta}_i^*, \beta_i)|Y=y\right) \ge 0.5 \ \text{ for any } y \in \mathcal{Y}.$$

i.e., $\hat{\beta}_{iB}$ is superior over $\hat{\beta}_i^{\,*}$ under PPC criterion, where $\mathcal{Y}$ is sample space.

Proof: We only prove the case of $i = 1$. The proof of $\hat{\beta}_{2B}$ is similar. Note that

$$\begin{aligned} W\left(\hat{\beta}_{1B}, \hat{\beta}_1^*; \beta_1\right) &= ||\hat{\beta}_{1B} - \beta_1||^2 - ||\hat{\beta}_1^* - \beta_1||^2 \\ &= \left[(\hat{\beta}_{1B} - \hat{\beta}_1^*) + (\hat{\beta}_1^* - \beta_1)\right]' \left[(\hat{\beta}_{1B} - \hat{\beta}_1^*) + (\hat{\beta}_1^* - \beta_1)\right] - ||\hat{\beta}_1^* - \beta_1||^2 \\ &= ||\hat{\beta}_1^* - \hat{\beta}_{1B}||^2 + 2(\hat{\beta}_1^* - \hat{\beta}_{1B}) \left[(\beta_1 - \hat{\beta}_{1B}) - (\hat{\beta}_1^* - \hat{\beta}_{1B})\right] \\ &= 2(\hat{\beta}_1^* - \hat{\beta}_{1B})'(\beta_1 - \hat{\beta}_{1B}) - ||\hat{\beta}_1^* - \hat{\beta}_{1B}||^2. \end{aligned} \tag{4.13}$$

Obviously

$$W \leq 0 \Longleftrightarrow 2(\hat{\beta}_1^* - \hat{\beta}_{1B})'(\beta_1 - \hat{\beta}_{1B}) \leq ||\hat{\beta}_1^* - \hat{\beta}_{1B}||^2.$$

From (4.12) we know that the prior of β_1 is normal distribution, hence the posterior distribution of β_1 given $Y = y$ is still normal distribution. Under the quadratic loss function the Bayes estimator of β_1 is $E(\beta_1|y)$, which has the same expression as (2.10), therefore the posterior of $\beta_1 - \hat{\beta}_{1B} = \beta_1 - E(\beta_1|y)$ given $Y = y$ is distributed as p_1-dimension normal distribution with zero mean. Thus, the posterior of $2(\hat{\beta}_1^* - \hat{\beta}_{1B})'(\beta_1 - \hat{\beta}_{1B})$ given $Y = y$ is one dimension normal distribution with zero mean. Hence we have

$$\begin{aligned} P_\pi\left(L(\hat{\beta}_{1B}, \beta_1)) \leq L(\hat{\beta}_1{}^*, \beta_1)|y\right) &= P_\pi\left(W\left(\hat{\beta}_{1B}, \hat{\beta}_1^*; \beta_1\right) \leq 0|y\right) \\ &= P_\pi\left(2(\hat{\beta}_1^* - \hat{\beta}_{1B})'(\beta_1 - \hat{\beta}_{1B}) \leq ||\hat{\beta}_1^* - \hat{\beta}_{1B}||^2|y\right) \\ &> P_\pi\left(2(\hat{\beta}_1^* - \hat{\beta}_{1B})'(\beta_1 - \hat{\beta}_{1B}) \leq 0 \mid y\right) = 0.5, \end{aligned}$$

the last inequality is true due to $||\hat{\beta}_1^* - \hat{\beta}_{1B}||^2 > 0$ with probability one. The Theorem 4.3 has been proved.

5 The MBRLU Estimators of Estimatable Function and Its Superiorities

In this section we consider the case that in (1.2) the rank of X is non-full rank, i.e., $R(X) < p$. Note that in this case β is un-estimable, hence we consider the estimable function $\eta = X\beta$. Since any estimable function of β can be expressed by the linear function of η, it is enough to study the estimator of η. It is obvious that the GLS estimator of $\eta = X\beta$ is

$$\hat{\eta}_{GLS} = X(X^{'}\Phi^{-1}X)^{-}X^{'}\Phi^{-1}Y,$$

where the expression of $\widehat{\eta}_{GLS}$ does not depend on g-inverse $(X^{'}\Phi^{-1}X)^{-}$, so it can be replaced by Moore-Penrose inverse $(X^{'}\Phi^{-1}X)^{+}$. Then we have

$$\hat{\eta}_{GLS} = X(X^{'}\Phi^{-1}X)^{+}X^{'}\Phi^{-1}Y = \begin{pmatrix} \hat{\eta}_1^* \\ \hat{\eta}_2^* \end{pmatrix}, \tag{5.1}$$

where η_i^* is the GLS estimator of $\eta_i = X_i\beta_i$ (i=1,2), and

$$\begin{aligned} \hat{\eta}_1^* &= X_1(X_1^{'}X_1)^{+}X_1^{'}Y_1 - \frac{\sigma_{12}}{\sigma_{22}}X_1(X_1^{'}X_1)^{+}X_1^{'}Y_2 \\ &= \hat{\eta}_1 - \frac{\sigma_{12}}{\sigma_{22}}X_1(X_1^{'}X_1)^{+}X_1^{'}Y_2, \end{aligned} \tag{5.2}$$

$$\begin{aligned} \hat{\eta}_2^* &= X_2(X_2^{'}X_2)^{+}X_2^{'}Y_2 - \frac{\sigma_{12}}{\sigma_{11}}X_2(X_2^{'}X_2)^{+}X_2^{'}Y_1 \\ &= \hat{\eta}_2 - \frac{\sigma_{12}}{\sigma_{11}}X_2(X_2^{'}X_2)^{+}X_2^{'}Y_1, \end{aligned} \tag{5.3}$$

with $\hat{\eta}_1 = X_1(X_1^{'}X_1)^{+}X_1^{'}Y_1,\ \hat{\eta}_2 = X_2(X_2^{'}X_2)^{+}X_2^{'}Y_2$.

Similar to the way used in section 1, we may obtain the MBRLU estimator of $\eta = X\beta$. Let $\eta_0 = X\mu$. Then we have

$$\hat{\eta}_B = XVX^{'}(XVX^{'} + \Phi)^{-1}Y + [I - XVX^{'}(XVX^{'} + \Phi)^{-1}]\ \eta_0. \tag{5.4}$$

By (2.6) and the fact that $X^{'}\Phi^{-1} = X^{'}\Phi^{-1}X(X^{'}\Phi^{-1}X)^{+}X^{'}\Phi^{-1}$ we know that

$$\begin{aligned} &XVX^{'}(XVX^{'} + \Phi)^{-1} = XVX^{'}[\Phi^{-1} - \Phi^{-1}X(V^{-1} + X^{'}\Phi^{-1}X)^{-1}X^{'}\Phi^{-1}] \\ &= XV[I - X^{'}\Phi^{-1}X(V^{-1} + X^{'}\Phi^{-1}X)^{-1}]X^{'}\Phi^{-1} \\ &= X(V^{-1} + X^{'}\Phi^{-1}X)^{-1}X^{'}\Phi^{-1} = H \\ &= X(V^{-1} + X^{'}\Phi^{-1}X)^{-1}X^{'}\Phi^{-1}X(X^{'}\Phi^{-1}X)^{+}X^{'}\Phi^{-1} \\ &= HX(X^{'}\Phi^{-1}X)^{+}X^{'}\Phi^{-1}. \end{aligned} \tag{5.5}$$

Therefore, using (5.4) and (5.5) and the fact that $X(X^{'}X)^{+}X^{'}X = X$ we have

$$\begin{aligned}
\hat{\eta}_B &= HX(X^{'}\Phi^{-1}X)^{+}X^{'}\Phi^{-1}Y + (I-H)\eta_0 \\
&= H\hat{\eta}_{GLS} + (I-H)\eta_0 = \hat{\eta}_{GLS} - (I-H)(\hat{\eta}_{GLS} - \eta_0) \\
&= \hat{\eta}_{GLS} - [X - HX(X^{'}X)^{+}X^{'}X](\hat{\beta}_{GLS} - \mu) \\
&= \hat{\eta}_{GLS} - [I - HX(X^{'}X)^{+}X^{'}](\hat{\eta}_{GLS} - \eta_0). \qquad (5.6)
\end{aligned}$$

By the assumption $X_1'X_2 = 0$ we have

$$\begin{aligned}
&X^{'}\Phi^{-1}X(X^{'}X)^{+}X^{'} \\
&= \begin{pmatrix} X_1^{'} & 0 \\ 0 & X_2^{'} \end{pmatrix} \begin{pmatrix} \sigma_{11.2}^{-1}I_n & -\rho_{12}I_n \\ -\rho_{12}I_n & \sigma_{22.1}^{-1}I_n \end{pmatrix} \begin{pmatrix} X_1(X_1^{'}X_1)^{+}X_1^{'} & 0 \\ 0 & X_2(X_2^{'}X_2)^{+}X_2^{'} \end{pmatrix} \\
&= \begin{pmatrix} \sigma_{11.2}^{-1}X_1^{'}X_1(X_1^{'}X_1)^{+}X_1^{'} & 0 \\ 0 & \sigma_{22.1}^{-1}X_2^{'}X_2(X_2^{'}X_2)^{+}X_2^{'} \end{pmatrix} = \begin{pmatrix} \sigma_{11.2}^{-1}X_1^{'} & 0 \\ 0 & \sigma_{22.1}^{-1}X_2^{'} \end{pmatrix}, \qquad (5.7)
\end{aligned}$$

and

$$\begin{aligned}
&HX(X^{'}X)^{+}X^{'} = X[V^{-1} + X^{'}\Phi^{-1}X]^{-1}X^{'}\Phi^{-1}X(X^{'}X)^{+}X^{'} \\
&= \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma_{11.2}}X_1^{'}X_1 + \frac{1}{\tau_1^2}I_{p_1} & 0 \\ 0 & \frac{1}{\sigma_{22.1}}X_2^{'}X_2 + \frac{1}{\tau_2^2}I_{p_2} \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{11.2}^{-1}X_1^{'} & 0 \\ 0 & \sigma_{22.1}^{-1}X_2^{'} \end{pmatrix} \\
&= \begin{pmatrix} X_1\left(X_1^{'}X_1 + \delta_1 I_{p_1}\right)^{-1}X_1^{'} & 0 \\ 0 & X_2\left(X_2^{'}X_2 + \delta_2 I_{p_2}\right)^{-1}X_2^{'} \end{pmatrix} \qquad (5.8)
\end{aligned}$$

with $\delta_1 = \sigma_{11.2}\tau_1^{-2}, \quad \delta_2 = \sigma_{22.1}\tau_2^{-2}$.

Substituting (5.8) into (5.6) we obtain

$$\begin{aligned}
\hat{\eta}_B &= \hat{\eta}_{GLS} - [I - HX(X^{'}X)^{+}X^{'}](\hat{\eta}_{GLS} - \eta_0) \\
&= \begin{pmatrix} \hat{\eta}_1^{*} - [I_n - X_1(X_1^{'}X_1 + \delta_1 I_{p_1})^{-1}X_1^{'}](\hat{\eta}_1^{*} - \eta_{01}) \\ \hat{\eta}_2^{*} - [I_n - X_2(X_2^{'}X_2 + \delta_2 I_{p_2})^{-1}X_2^{'}](\hat{\eta}_2^{*} - \eta_{02}) \end{pmatrix}, \qquad (5.9)
\end{aligned}$$

where $\eta_0' = (\eta_{01}', \eta_{02}')$ and $\eta' = (\eta_1', \eta_2')$. So the MBRLU estimators of $\eta_i (i = 1, 2)$ are

$$\hat{\eta}_{1B} = \hat{\eta}_1^{*} - [I_n - X_1(X_1^{'}X_1 + \delta_1 I_{p_1})^{-1}X_1^{'}]\left(\hat{\eta}_1^{*} - \eta_{01}\right), \qquad (5.10)$$

$$\hat{\eta}_{2B} = \hat{\eta}_2^{*} - [I_n - X_2(X_2^{'}X_2 + \delta_2 I_{p_2})^{-1}X_2^{'}]\left(\hat{\eta}_2^{*} - \eta_{02}\right). \qquad (5.11)$$

Theorem 5.1 Let the GLS estimators and MBRLU estimators of $\eta_i = X_i\beta_i$ are given by (5.2), (5.3) and (5.10),(5.11), respectively, then

$$M(\hat{\eta}_i^*) - M\left(\hat{\eta}_{iB}\right) \geq 0, \quad \text{i=1, 2.}$$

Proof: We only prove the case of $i = 1$, the case for $i = 2$ can be proved in a similar way. Let $G_1 = I_n - X_1(X_1'X_1 + \delta_1 I_{p_1})^{-1}X_1'$, then

$$\begin{aligned}
M\left(\hat{\eta}_{1B}\right) &= E\left\{[\hat{\eta}_1^* - G_1\left(\hat{\eta}_1^* - \eta_{01}\right) - \eta_1][\hat{\eta}_1^* - G_1\left(\hat{\eta}_1^* - \eta_{01}\right) - \eta_1]'\right\} \\
&= E\left\{[(\hat{\eta}_1^* - \eta_1) - G_1(\hat{\eta}_1^* - \eta_{01})][(\hat{\eta}_1^* - \eta_1) - G_1(\hat{\eta}_1^* - \eta_{01})]'\right\} \\
&= E\left\{\left(\hat{\eta}_1^* - \eta_1\right)\left(\hat{\eta}_1^* - \eta_1\right)' - \left(\hat{\eta}_1^* - \eta_1\right)\left(\hat{\eta}_1^* - \eta_{01}\right)' G_1' \right. \\
&\quad \left. - G_1\left(\hat{\eta}_1^* - \eta_{01}\right)\left(\hat{\eta}_1^* - \eta_1\right)' + G_1\left(\hat{\eta}_1^* - \eta_{01}\right)\left(\hat{\eta}_1^* - \eta_{01}\right)' G_1'\right\} \\
&= M(\hat{\eta}_1^*) - K_1G_1' - G_1K_1' + G_1K_2G_1', \qquad (5.12)
\end{aligned}$$

where

$$\begin{aligned}
K_2 &= E\left(\left(\hat{\eta}_1^* - \eta_{01}\right)\left(\hat{\eta}_1^* - \eta_{01}\right)'\right) = Cov\left(\hat{\eta}_1^*\right) \\
&= Cov[E\left(\hat{\eta}_1^*|\eta\right)] + E[Cov\left(\hat{\eta}_1^*|\eta\right)] \\
&= \tau_1^2 X_1X_1' + \sigma_{11.2}X_1(X_1'X_1)^+X_1', \qquad (5.13) \\
K_1 &= E[\left(\hat{\eta}_1^* - \eta_1\right)\left(\hat{\eta}_1^* - \eta_{01}\right)'] \\
&= Cov\left(\hat{\eta}_1^*\right) - Cov(\eta_1) = \sigma_{11.2}X_1(X_1'X_1)^+X_1'. \qquad (5.14)
\end{aligned}$$

Putting (5.13), (5.14) into (5.12) and by the fact of $G_1 = I_n - X_1(X_1'X_1 + \delta_1 I_{p_1})^{-1}X_1' = (I_n + \delta_1^{-1}X_1X_1')^{-1}$, we have

$$\begin{aligned}
M\left(\hat{\eta}_1^*\right) - M\left(\hat{\eta}_{1B}\right) &= K_1G_1' + G_1K_1' - G_1K_2G_1' \\
&= G_1\left(\tau_1^2 X_1X_1' + \sigma_{11.2}X_1(X_1'X_1)^+X_1'\right)G_1' \geq 0.
\end{aligned}$$

Theorem 5.1 has been proved. □

For nay general estimable functions $\gamma_i = P_i\beta_i$, where P_i is a $k_i \times p_i$ real matrix, for which there exists $k_i \times n$ matrix C_i such that $P_i = C_iX_i$, therefore,

$\gamma_i = C_i X_i \beta_i = C_i \eta_i (i = 1, 2)$. It is easy to know the GLS estimators and MBRLU estimators of γ_i would be

$$\hat{\gamma}_i^* = C_i \hat{\eta}_i^*, \qquad \hat{\gamma}_{iB} = C_i \hat{\eta}_{iB}, \qquad i = 1, 2, \tag{5.15}$$

then we have the following corollary.

Corollary 5.1 Let the GLS estimators and MBRLU estimators of estimable function γ_i $(i = 1, 2)$ are given by (5.15), then

$$M(\hat{\gamma}_i^*) - M(\hat{\gamma}_{iB}) \geq 0, \quad i = 1,\ 2.$$

Proof: The conclusion holds since

$$M(\hat{\gamma}_i^*) - M(\hat{\gamma}_{iB}) = C_i \left[M(\hat{\eta}_i^*) - M\left(\hat{\eta}_{iB}\right) \right] C_i^{'} \geq 0.$$

Now we discuss the superiority of MBRLU estimator of η_1 under PRPC criterion, the superiority of MBRLU estimator of η_2 can be discussed similarly. Under the loss function (2.2) we have the following result:

Theorem 5.2 Let the GLS estimator and MBRLU estimator of $\eta_1 = X_1 \beta_1$ be given by (5.2) and (5.10). Under the condition (4.1) and suppose $R(X_1) = t_1 \leq p_1$, if

$$\frac{\tau_1^2}{\sigma_{11.2}} \leq \frac{\tilde{\lambda}_{t_1}(t_1 - 2)}{2 t_1 \tilde{\lambda}_1^2} \tag{5.16}$$

then we have

$$P_\pi \left(L(\hat{\eta}_{1B} - \eta_1)) \leq L(\hat{\eta_1}^* - \eta_1) \right) \geq 0.5, \quad \text{for every } \pi \in \Gamma(\beta_1),$$

where $\tilde{\lambda}_1, \tilde{\lambda}_{t_1}$ are the maximum and minimum positive eigenvalue of $X_1^{'} X_1$, and $\Gamma(\beta_1)$ is given by Theorem 4.1.

Remark 5.1 The condition of (5.16) indicates the fact that the variance of prior should not be too larger than that of the samples. It implies some requirement for the precision of the variance of prior distribution.

Proof: Let $G_1 = I - X_1(X_1'X_1 + \delta_1 I_{p_1})^{-1}X_1' = (I + \delta_1^{-1}X_1X_1')^{-1}$ and $W(\hat{\eta}_{1B}, \eta_1^*; \eta_1) = L(\hat{\eta}_{1B}, \eta_1) - L(\hat{\eta}_1{}^*, \eta_1)$, where δ_1 is given in (5.8). By (5.10) we have

$$\begin{aligned} L(\hat{\eta}_{1B}, \eta_1) &= [(\hat{\eta}_1^* - \eta_1) - G_1(\hat{\eta}_1^* - \eta_{01})]'[(\hat{\eta}_1^* - \eta_1) - G_1(\hat{\eta}_1^* - \eta_{01})] \\ &= L(\hat{\eta}_1^*, \eta_1) - 2(\hat{\eta}_1^* - \eta_{01})'G_1(\hat{\eta}_1^* - \eta_1) \\ &\quad + (\hat{\eta}_1^* - \eta_{01})'G_1^2(\hat{\eta}_1^* - \eta_{01}). \end{aligned} \tag{5.17}$$

From (5.17), it is easy to see that $W < 0$ is equivalent to

$$(\hat{\eta}_1^* - \eta_{01})'G_1^2(\hat{\eta}_1^* - \eta_{01}) \leq 2(\hat{\eta}_1^* - \eta_{01})'G_1(\hat{\eta}_1^* - \eta_1). \tag{5.18}$$

since $G_1^2 < G_1$, then (5.18) is implied by

$$(\hat{\eta}_1^* - \eta_{01})'G_1(\hat{\eta}_1^* - \eta_{01}) \leq 2(\hat{\eta}_1^* - \eta_{01})'G_1(\hat{\eta}_1^* - \eta_1). \tag{5.19}$$

Substituting $\hat{\eta}_1^* - \eta_{01} = \hat{\eta}_1^* - \eta_1 - (\eta_{01} - \eta_1)$ into (5.19), we have

$$(\eta_1 - \eta_{01})'G_1(\eta_1 - \eta_{01}) \leq (\hat{\eta}_1^* - \eta_1)'G_1(\hat{\eta}_1^* - \eta_1). \tag{5.20}$$

Since $\tilde{\lambda}_1$ and $\tilde{\lambda}_{t_1}$ is the maximum and minimum non-zero eigenvalue of X_1X_1' respectively, then

$$\left(\frac{\tau_1^2}{\sigma_{11.2}}\tilde{\lambda}_1 + 1\right)^{-1} I_n \leq G_1 = \left(\frac{\tau_1^2}{\sigma_{11.2}}X_1X_1' + I_n\right)^{-1} \leq \left(\frac{\tau_1^2}{\sigma_{11.2}}\tilde{\lambda}_{t_1} + 1\right)^{-1} I_n.$$

Hence (5.20) is implied by

$$\left(\eta_1^{-1}\tilde{\lambda}_{t_1} + 1\right)^{-1}(\eta_1 - \eta_{01})'(\eta_1 - \eta_{01}) \leq \left(\delta_1^{-1}\tilde{\lambda}_1 + 1\right)^{-1}(\hat{\eta}_1^* - \eta_1)'(\hat{\eta}_1^* - \eta_1). \tag{5.21}$$

Note that $0 < \frac{\delta_1^{-1}\tilde{\lambda}_1 + 1}{\delta_1^{-1}\tilde{\lambda}_{t_1} + 1} \leq \frac{\tilde{\lambda}_1}{\tilde{\lambda}_{t_1}}$, (5.21) is implied by

$$\frac{\tilde{\lambda}_1}{\tilde{\lambda}_{t_1}}(\eta_1 - \eta_{01})'(\eta_1 - \eta_{01}) \leq (\hat{\eta}_1^* - \eta_1)'(\hat{\eta}_1^* - \eta_1). \tag{5.22}$$

From (4.1) we know that

$$(\hat{\eta}_1^* - \eta_1)|\eta \sim N\left(0, \sigma_{11.2}X_1(X_1'X_1)^+X_1'\right).$$

Since $Q = X_1(X_1'X_1)^+X_1'$ is an idempotent matrix, there exists an orthogonal matrix P such that

$$PQP' = \begin{pmatrix} I_{t_1} & 0 \\ 0 & 0 \end{pmatrix}.$$

Let $\tilde{Z} = \sigma_{11.2}^{-1/2} P \left(\hat{\eta}_1^* - \eta_1\right)$, then

$$\tilde{Z}|\delta_1 \sim N\left(0, \begin{pmatrix} I_{t_1} & 0 \\ 0 & 0 \end{pmatrix}\right),$$

which implies $\tilde{Z}'\tilde{Z} \sim \chi^2_{t_1}$. Thus (5.22) is equivalent to

$$\frac{\tilde{\lambda}_1}{\tilde{\lambda}_{t_1}\sigma_{11.2}} \left|\left|\eta_1 - \eta_{01}\right|\right|^2 \leq \tilde{Z}'\tilde{Z}. \tag{5.23}$$

by (5.18)-(5.23) and Markov-inequality we obtain

$$\begin{aligned}
&P_\pi\left(W\left(\hat{\eta}_{1B}, \eta_1^*; \eta_1\right) \leq 0\right) \geq P_\pi\left(\tilde{Z}'\tilde{Z} \geq \frac{\tilde{\lambda}_1}{\sigma_{11.2}\tilde{\lambda}_{t_1}} \left|\left|\eta_1 - \eta_{01}\right|\right|^2\right) \\
&= 1 - P_\pi\left(\left|\left|\eta_1 - \eta_{01}\right|\right|^2 \geq \frac{\sigma_{11.2}\tilde{\lambda}_{t_1}}{\tilde{\lambda}_1}\tilde{Z}'\tilde{Z}\right) \\
&= 1 - E\left[P_\pi\left(\left|\left|\eta_1 - \eta_{01}\right|\right|^2 \geq \frac{\sigma_{11.2}\tilde{\lambda}_{t_1}}{\tilde{\lambda}_1}\tilde{Z}'\tilde{Z}\,\middle|\,\tilde{Z}'\tilde{Z}\right)\right] \\
&\geq 1 - \frac{\tilde{\lambda}_1 E\left|\left|\eta_1 - \eta_{01}\right|\right|^2}{\sigma_{11.2}\tilde{\lambda}_{t_1}} E\left(\frac{1}{\tilde{Z}'\tilde{Z}}\right) = 1 - \frac{\tilde{\lambda}_1 tr(Cov(\eta_1))}{\sigma_{11.2}\tilde{\lambda}_{t_1}(t_1 - 2)} \\
&\geq 1 - \frac{\tilde{\lambda}_1 \tau_1^2 tr(X_1'X_1)}{\sigma_{11.2}\tilde{\lambda}_{t_1}(t_1 - 2)} \geq 1 - \frac{t_1\tilde{\lambda}_1^2\tau_1^2}{\sigma_{11.2}\tilde{\lambda}_{t_1}(t_1 - 2)} \geq 0.5.
\end{aligned}$$

The proof of Theorem 5.2 is completed. □

Similar to Theorem 5.2, we have the following theorem.

Theorem 5.3 Let the GLS estimator and MBRLU estimator of $\eta_2 = X_2\beta_2$ be given by (5.3) and (5.11). Under the condition (4.1) and suppose $R(X_2) = t_2 \leq p_2$, if

$$\frac{\tau_2^2}{\sigma_{22.1}} \leq \frac{\bar{\lambda}_{t_2}(t_2 - 2)}{2t_2\bar{\lambda}_1^2} \tag{5.24}$$

then we have

$$P_\pi\left(L(\hat{\eta}_{2B}-\eta_2)\right)\leq L(\hat{\eta_2}^*-\eta_2))\geq 0.5, \quad \text{for every } \pi\in\Gamma(\beta_2),$$

where $\bar{\lambda}_1, \bar{\lambda}_{t_2}$ are the maximum and minimum positive eigenvalue of $X_2'X_2$, and $\Gamma(\beta_2)$ is the same as that of Theorem 4.2.

Similar to the proof of Theorem 4.3, we have the following result.

Theorem 5.4 Let the GLS estimator and MBRLU estimator of $\eta_i = X_i\beta_i$ be given by (5.2), (5.3) and (5.10), (5.11) respectively. Under the conditions (4.1) and (4.12). Then for $i = 1, 2$, we have

$$P_\pi\left(L(\hat{\eta}_{iB},\eta_i)\right)\leq L(\hat{\eta_i}^*,\eta_i)|y)\geq 0.5, \quad \text{for any } y\in\mathcal{Y},$$

where $\mathcal{Y}$ is the sample space.

6 Concluding remarks

In summary, we have investigated Bayesian estimation problem of regression parameter in the system of two seemingly unrelated regressions. We derive the Bayes minimum risk linear unbiased (MBRLU) estimators for regression parameters and establish their superiorities based on the mean square error matrix (MSEM) criterion. Also, we exhibit the superiorities of MBRLU estimators in terms of the predictive Pitman closeness (PRPC) criterion and the posterior Pitman closeness (PPC) criterion, respectively. In the case that the design matrices are non-full rank, the superiorities of BMRLU estimators of some estimable functions are investigated.

References

[1] J.O. Berger, *Statistical decision theory and Bayesian analysis*, Berlin, Springer, 1985.

[2] G.E.P. Box and G. Tiao, *Bayesian Inference in Statistical Analysis*, Reading, Addison-Wesley, 1973.

[3] R.L. Fountain and J.P. Keating, The Pitman comparison of unbiased linear estimators, *Statistics and Probability Letters*, **19**, (1994), 131-136.

[4] M. Ghosh and P.K. Sen, Bayesian Pitman closeness, *Comm. Statist. Theory and Methods*, **20**, (1991),3659-3678.

[5] M.H.J.Gruber, *Regression Estimators, A Comparative Study*, Boston, Academic Press, 1990.

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

[7] J.P. Keating, R.L. Mason and P.K. Sen, Pitman's Measure of Closeness: A Comparison of Statistical Estimators, Philadelphia-Society of Industrial and Applied Mathematics, (1993).

[8] J. Kementa and R.F.Gilbert, Small sample properties of alternative estimators of seemingly unrelated regressions, *J. Amer. Statist. Assoc.*, **63**, (1968), 1180-1200.

[9] C.T. Lin, The efficiency of least squares estimator of a seemingly unrelated regression model, *Comm. Statist. Simulation*, **20**, (1991), 919-925.

[10] A.Y. Liu, Efficient estimation of two seemingly unrelated regression equations, *Journal of Multivariate Analysis*, **82**, (2002), 445-456.

[11] R.L. Mason, J.P. Keating, P.K. Sen and N.W.Blaylock, Comparison of linear estimators using Pitman's measure of closeness, *J. Amer. Statist. Assoc.*, **81**, (1990), 579-581.

[12] J.S. Mehta and P.A.V.B. Swamy, Further evidence on the relative efficiencies of Zellner's seemingly unrelated regressions estimator, *J. Amer. Statist. Assoc.*, **71**, (1976), 634-639.

[13] E.J.G. Pitman, The closest estimates of statistical parameters, *Proc. Camb. Phil. Soc.*, **33**, (1937), 212-222.

[14] C.R. Rao, *Linear Statistical Inference and Its Applications*, Second Edition, New York, Wiley, 1973.

[15] C.R. Rao, Estimation of parameters in a linear model, The 1975 Wald Memorial Lectures, *Ann. Statist.*, **4**, (1976), 1023-1037.

[16] C.R. Rao, *Some comments on the minimum mean square as criterion of estimation. In Statistics and Related Topics*, North Holland, Amsterdan, pp. 123-143, 1981.

[17] C.R. Rao, J.P. Keating and R.L. Mason, The Pitman nearness criterion and determination, *Comm. Statist. Theor. Math.*, **15**, (1986), 3173-3191.

[18] N.S. Revankar, Some finite sample results in the context of two seemingly unrelated regression equations, *J. Amer. Statist. Assoc.*, **69**, (1974), 187-190.

[19] P. Schmidt, Estimation of seemingly unrelated regressions with unequal numbers of observations, *Journal of Econometrics*, **5**, (1977), 365-377.

[20] G. Trenkler and L.S. Wei, The Bayes estimators in a misspecified linear regression model, *Test*, **5**, (1996), 113-123.

[21] L.C. Wang and N. Veraverbeke, MSE superiority of Bayes and empirical Bayes estimators in two generalized seemingly unrelated regressions, *Statistics & Probability Letters*, **78**, (2008), 109-117.

[22] L.C. Wang, H. Lian and R.S. Singh, On efficient estimators of two seemingly unrelated regressions, *Statistics & Probability Letters*, **81**, (2011), 563-570.

[23] S.G. Wang, A new estimate of regression coefficients in seemingly unrelated regression system, *Science in China* (Ser. A), **7**, (1989), 808-816.

[24] S.G. Wang and S.C. Chow, *Advanced linear models: Theory and Applications*, New York, Marcel Dekker, 1994.

[25] W.P. Zhang, *Research of Problems of Bayes Analysis in Linear Model*, Thesis of USTC for Ph.D, 2005-JX-C-1773, 2005.

[26] A. Zellner, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *J. Amer. Statist. Assoc.*, **57**, (1962), 348-368.

[27] A. Zellner, Estimators of seemingly unrelated regressions: some exact finite sample results, *J. Amer. Statist. Assoc.*, **58**, (1963), 977-992.