The equalities of ordinary least-squares estimators and best linear unbiased estimators for the restricted linear model

Yongge Tian[a] and Douglas P. Wiens[b]

[a]School of Economics, Shanghai University of Finance and Economics, Shanghai 200433, China

[b]Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1

Abstract. We investigate in this paper a variety of equalities for the ordinary least-squares estimators and the best linear unbiased estimators under the general linear (Gauss-Markov) model $\{\mathbf{y},\, \mathbf{X}\boldsymbol{\beta},\, \sigma^2\boldsymbol{\Sigma}\}$ and the restrained model $\{\mathbf{y},\, \mathbf{X}\boldsymbol{\beta} \,|\, \mathbf{A}\boldsymbol{\beta} = \mathbf{b},\, \sigma^2\boldsymbol{\Sigma}\}$.

Keywords: OLSE; BLUE; general linear model; linear restriction; parametric function; matrix rank method

1 Introduction

Suppose we are given the following general linear (Gauss-Markov) model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad E(\boldsymbol{\varepsilon}) = \mathbf{0} \ \text{ and } \ Cov(\boldsymbol{\varepsilon}) = \sigma^2\boldsymbol{\Sigma}, \tag{1}$$

constrained by a consistent linear matrix equation

$$\mathbf{A}\boldsymbol{\beta} = \mathbf{b}, \tag{2}$$

where $\mathbf{X}$ is a known $n \times p$ matrix with $\mathrm{rank}(\mathbf{X}) \leq p$, $\mathbf{A}$ is a known $m \times p$ matrix with $\mathrm{rank}(\mathbf{A}) = m$, $\mathbf{b}$ is a known $m \times 1$ vector, $\mathbf{y}$ is an $n \times 1$ observable random vector, $\boldsymbol{\beta}$ is a $p \times 1$ vector of parameters to be estimated, $0 < \sigma^2$ is an unknown parameter, and $\boldsymbol{\Sigma}$ is an $n \times n$ known nonnegative definite matrix. The model (1) is often written as the following triplet

$$\mathscr{M} = \{\mathbf{y},\, \mathbf{X}\boldsymbol{\beta},\, \sigma^2\boldsymbol{\Sigma}\}, \tag{3}$$

the model (1) with the parameters constrained by the equation (2) is written as

$$\mathscr{M}_c = \{\mathbf{y},\, \mathbf{X}\boldsymbol{\beta} \,|\, \mathbf{A}\boldsymbol{\beta} = \mathbf{b},\, \sigma^2\boldsymbol{\Sigma}\}. \tag{4}$$

In the investigation of linear models, parameter constraints are usually handled by transforming an explicitly constrained model into an implicitly constrained model. The most popular transformations are based on model reduction, Lagrangian multipliers, and reparameterization by general solution of the equation (2). Through block matrices, (1) and (2) can be written as

$$\mathscr{M}_c = \left\{ \begin{bmatrix} \mathbf{y} \\ \mathbf{b} \end{bmatrix},\ \begin{bmatrix} \mathbf{X} \\ \mathbf{A} \end{bmatrix}\boldsymbol{\beta},\ \sigma^2\begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right\}. \tag{5}$$

If the observable vector $\mathbf{y}$ in (3) arises from the combination of two sub-samples $\mathbf{y}_1$ and $\mathbf{y}_2$, where $\mathbf{y}_1$ and $\mathbf{y}_2$ are $n_1 \times 1$ and $n_2 \times 1$ vectors with $n_1 + n_2 = n$, then (3) can also be written as a partitioned form

$$\mathscr{M} = \left\{ \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix},\ \begin{bmatrix} \mathbf{X}_{(1)} \\ \mathbf{X}_{(2)} \end{bmatrix}\boldsymbol{\beta},\ \sigma^2\begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \right\}, \tag{6}$$

E-mail addresses: yongge@mail.shufe.edu.cn, doug.wiens@ualberta.ca

where $\mathbf{X}_{(1)}$ and $\mathbf{X}_{(2)}$ are $n_1 \times p$, $n_2 \times p$ matrices, $\mathbf{\Sigma}_{11}$ is an $n_1 \times n_1$ matrix. In this case, the first sub-sample model in (6) can be written as

$$\mathscr{M}_1 = \{\mathbf{y}_1, \mathbf{X}_{(1)}\boldsymbol{\beta},\ \sigma^2\mathbf{\Sigma}_{11}\}, \tag{7}$$

and the first sub-sample model in (4) can be written as

$$\mathscr{M}_{1c} = \{\mathbf{y}_1,\ \mathbf{X}_{(1)}\boldsymbol{\beta} \mid \mathbf{A}\boldsymbol{\beta} = \mathbf{b},\ \sigma^2\mathbf{\Sigma}_{11}\}. \tag{8}$$

Because (4) has a linear restriction on $\boldsymbol{\beta}$, the estimation of $\boldsymbol{\beta}$ is more complicated than that of $\boldsymbol{\beta}$ under (3). Of interest to us is to consider relations between estimators of $\boldsymbol{\beta}$ under (3) and (4). In particular, it is important to give necessary and sufficient conditions for the estimators of $\boldsymbol{\beta}$ under (3) and (4) to be equal. If two estimators of $\boldsymbol{\beta}$ under (3) and (4) are equal, it is natural to use the estimator of $\boldsymbol{\beta}$ under (3) instead of that of (4). The restricted linear model (4) also has a closed link with the hypothesis test $\mathbf{A}\boldsymbol{\beta} = \mathbf{b}$ for $\boldsymbol{\beta}$ under (3). From the equalities of the estimators of $\boldsymbol{\beta}$ under (3) and (4), one can also check the testability of the hypothesis. Linear models with restrictions occur widely in statistics, see, e.g., Amemiya (1985), Bewley (1986), Chipman and Rao (1964), Dent (1980), Haupt and Oberhofer (2002), Ravikumar, Ray and Savin (2000) and Theil (1971).

Throughout this paper, $\mathbb{R}^{m\times n}$ stands for the set of all real $m \times n$ matrices; $\mathbf{M}'$, $r(\mathbf{M})$ and $\mathscr{R}(\mathbf{M})$ denote the transpose, the rank and the range (column space) of a matrix $\mathbf{M}$, respectively. For $\mathbf{M} \in \mathbb{R}^{m\times n}$, $\mathbf{M}^-$ denotes a g-inverse of $\mathbf{M}$, i.e., $\mathbf{M}\mathbf{M}^-\mathbf{M} = \mathbf{M}$; $\mathbf{M}^+$ denotes the Moore-Penrose inverse of $\mathbf{M}$, i.e., $\mathbf{M}\mathbf{M}^+\mathbf{M} = \mathbf{M}$, $\mathbf{M}^+\mathbf{M}\mathbf{M}^+ = \mathbf{M}^+$, $(\mathbf{M}\mathbf{M}^+)' = \mathbf{M}\mathbf{M}^+$ and $(\mathbf{M}^+\mathbf{M})' = \mathbf{M}^+\mathbf{M}$. Further, let $\mathbf{P}_\mathbf{M} = \mathbf{M}\mathbf{M}^+$, $\mathbf{E}_\mathbf{M} = \mathbf{I}_m - \mathbf{M}\mathbf{M}^+$ and $\mathbf{F}_\mathbf{M} = \mathbf{I}_n - \mathbf{M}^+\mathbf{M}$.

Let $\mathscr{M}$ be as given in (3). The OLSE of the parameter vector $\boldsymbol{\beta}$ under (3) is defined to be any vector minimizing $\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 = (\,\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\,)'(\,\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\,)$ and is denoted by $\widehat{\boldsymbol{\beta}}$, The OLSE of the parameter vector $\mathbf{X}\boldsymbol{\beta}$ under (3) is defined to be $\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}\widehat{\boldsymbol{\beta}}$. In general, for any $l \times p$ given matrix $\mathbf{K}$, the OLSE of the vector $\mathbf{K}\boldsymbol{\beta}$ under (3) is defined to be $\mathrm{OLSE}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{K}\widehat{\boldsymbol{\beta}}$. The BLUE of $\mathbf{X}\boldsymbol{\beta}$ under (3) is defined to be a linear estimator $\mathbf{G}\mathbf{y}$ such that $E(\mathbf{G}\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$ and for any other linear unbiased estimator $\mathbf{M}\mathbf{y}$ of $\mathbf{X}\boldsymbol{\beta}$, $Cov(\mathbf{M}\mathbf{y}) - Cov(\mathbf{G}\mathbf{y}) = \mathbf{M}\mathbf{\Sigma}\mathbf{M}' - \mathbf{G}\mathbf{\Sigma}\mathbf{G}'$ is nonnegative definite. The two estimators are well known and have been extensively investigated in the literature.

Some frequently used OLSEs and BLUEs under the models (3), (4), (6), (7) and (8) are

$$\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}),\ \ \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}),\ \ \mathrm{OLSE}_{\mathscr{M}}(\boldsymbol{\beta}),\ \ \mathrm{OLSE}_{\mathscr{M}_c}(\boldsymbol{\beta}),\ \ \mathrm{OLSE}_{\mathscr{M}_1}(\boldsymbol{\beta}),\ \ \mathrm{OLSE}_{\mathscr{M}_{1c}}(\boldsymbol{\beta}),$$

$$\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}),\ \ \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}),\ \ \mathrm{BLUE}_{\mathscr{M}}(\boldsymbol{\beta}),\ \ \mathrm{BLUE}_{\mathscr{M}_c}(\boldsymbol{\beta}),\ \ \mathrm{BLUE}_{\mathscr{M}_1}(\boldsymbol{\beta}),\ \ \mathrm{BLUE}_{\mathscr{M}_{1c}}(\boldsymbol{\beta}).$$

The OLSEs and BLUEs of $\mathbf{X}\boldsymbol{\beta}$ and $\boldsymbol{\beta}$ under (3) and (4) can be derived through matrix equations, generalized inverses of matrices or the Lagrange multipliers method.

A well-known result on the solvability and general solution of the linear matrix equation $\mathbf{A}\mathbf{x} = \mathbf{b}$ is given below.

Lemma 1 *The linear matrix equation* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *is consistent if and only if* $\mathbf{A}\mathbf{A}^-\mathbf{b} = \mathbf{b}$. *In this case, the general solution of* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *can be written as*

$$\mathbf{x} = \mathbf{A}^-\mathbf{b} + (\,\mathbf{I}_p - \mathbf{A}^-\mathbf{A}\,)\mathbf{v},$$

where the vector $\mathbf{v}$ *is arbitrary. If* $\mathbf{A}^-$ *is taken as the Moore-Penrose inverse* $\mathbf{A}^+$, *the general solution of* $\mathbf{A}\boldsymbol{\beta} = \mathbf{b}$ *can be written as*

$$\mathbf{x} = \mathbf{A}^+\mathbf{b} + (\,\mathbf{I}_p - \mathbf{A}^+\mathbf{A}\,)\mathbf{v}, \tag{9}$$

where the vector $\mathbf{v}$ *is arbitrary. If* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *is not consistent, then* (9) *is the general expression of the least-squares solutions of* $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Note that $[\mathbf{X}, \mathbf{\Sigma}][\mathbf{X}, \mathbf{\Sigma}]^{+}[\mathbf{X}, \mathbf{\Sigma}] = [\mathbf{X}, \mathbf{\Sigma}]$. Hence if the model in (1) is correct, then

$$E([\mathbf{X}, \mathbf{\Sigma}][\mathbf{X}, \mathbf{\Sigma}]^{+}\mathbf{y} - \mathbf{y}) = [\mathbf{X}, \mathbf{\Sigma}][\mathbf{X}, \mathbf{\Sigma}]^{+}\mathbf{X}\boldsymbol{\beta} - \mathbf{X}\boldsymbol{\beta} = \mathbf{0}$$

and

$$Cov([\mathbf{X}, \mathbf{\Sigma}][\mathbf{X}, \mathbf{\Sigma}]^{+}\mathbf{y} - \mathbf{y}) = \sigma^2([\mathbf{X}, \mathbf{\Sigma}][\mathbf{X}, \mathbf{\Sigma}]^{+} - \mathbf{I})\mathbf{\Sigma}([\mathbf{X}, \mathbf{\Sigma}][\mathbf{X}, \mathbf{\Sigma}]^{+} - \mathbf{I})' = \mathbf{0},$$

so that $[\mathbf{X}, \mathbf{\Sigma}][\mathbf{X}, \mathbf{\Sigma}]^{+}\mathbf{y} = \mathbf{y}$, or equivalently, $\mathbf{y} \in [\mathbf{X}, \mathbf{\Sigma}]$ holds with probability 1.

Definition 1 The linear model in (1) is said to be *consistent* if $\mathbf{y} \in [\mathbf{X}, \mathbf{\Sigma}]$ holds with probability 1.

The consistency concept for a general linear model was introduced in Rao (1971, 1973), which is a basic assumption in the investigation of the best linear unbiased estimator (BLUE) of $\mathbf{X}\boldsymbol{\beta}$ under the general linear model in (1).

Definition 2 Suppose $\mathscr{M}$ in (3) is consistent. Then two linear estimators $\mathbf{L}_1\mathbf{y}$ and $\mathbf{L}_2\mathbf{y}$ of $\mathbf{X}\boldsymbol{\beta}$ under (3) are said to be equal with probability 1 if $(\mathbf{L}_1 - \mathbf{L}_2)[\mathbf{X}, \mathbf{\Sigma}] = \mathbf{0}$.

Lemma 2 *Let $\mathscr{M}$ be as given in* (3), *and suppose $\mathbf{K}_1\mathbf{y} + \mathbf{z}_1$ and $\mathbf{K}_2\mathbf{y} + \mathbf{z}_2$ are two linear unbiased estimators for $\mathbf{X}\boldsymbol{\beta}$ under* (3). *Then $\mathbf{K}_1\mathbf{y} + \mathbf{z}_1 = \mathbf{K}_2\mathbf{y} + \mathbf{z}_2$ holds with probability 1 if and only if $\mathbf{K}_1\mathbf{\Sigma} = \mathbf{K}_2\mathbf{\Sigma}$ holds.*

Lemma 3 *Let $\mathscr{M}$ and $\mathscr{M}_c$ be as given in* (3) *and* (4). *Then:*

(a) *The OLSE of $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}$ can uniquely be written as*

$$\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{P}_{\mathbf{X}}\mathbf{y} \tag{10}$$

with $E[\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{X}\boldsymbol{\beta}$ and $Cov[\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})] = \sigma^2\mathbf{P}_{\mathbf{X}}\mathbf{\Sigma}\mathbf{P}_{\mathbf{X}}$.

(b) *The BLUE of $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}$ can be written as*

$$\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{y}, \tag{11}$$

where $\mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}$ is a solution of the consistent matrix equation $\mathbf{Z}[\mathbf{X}, \mathbf{\Sigma}\mathbf{E}_{\mathbf{X}}] = [\mathbf{X}, \mathbf{0}]$, where the two partitioned matrices satisfy

$$\mathscr{R}[\mathbf{X}, \mathbf{\Sigma}\mathbf{E}_{\mathbf{X}}] = \mathscr{R}[\mathbf{X}, \mathbf{\Sigma}] \ \ \textit{and} \ \ \mathscr{R}\begin{bmatrix} \mathbf{X}' \\ \mathbf{0} \end{bmatrix} \subseteq \mathscr{R}\begin{bmatrix} \mathbf{X}' \\ \mathbf{E}_{\mathbf{X}}\mathbf{\Sigma} \end{bmatrix}. \tag{12}$$

The general expression of $\mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}$ can be written as

$$\mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}} = [\mathbf{X}, \mathbf{0}][\mathbf{X}, \mathbf{\Sigma}\mathbf{E}_{\mathbf{X}}]^{+} + \mathbf{V}(\mathbf{I}_n - [\mathbf{X}, \mathbf{\Sigma}\mathbf{E}_{\mathbf{X}}][\mathbf{X}, \mathbf{\Sigma}\mathbf{E}_{\mathbf{X}}]^{+}), \tag{13}$$

where $\mathbf{V} \in \mathbb{R}^{n\times n}$ is arbitrary. Moreover,

$$E[\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{X}\boldsymbol{\beta} \ \ \textit{and} \ \ Cov[\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})] = \sigma^2\mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma}\mathbf{P}'_{\mathbf{X}\|\mathbf{\Sigma}}.$$

In particular, $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ *is unique with probability 1 if* (3) *is consistent.*

(c) *The OLSE of $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}_c$ can uniquely be written as*

$$\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) = (\mathbf{I}_n - \mathbf{P}_{\mathbf{X}_\mathbf{A}})\mathbf{X}\mathbf{A}^{+}\mathbf{b} + \mathbf{P}_{\mathbf{X}_\mathbf{A}}\mathbf{y}, \tag{14}$$

where $\mathbf{X}_\mathbf{A} = \mathbf{X}\mathbf{F}_\mathbf{A}$. In this case,

$$E[\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{X}\boldsymbol{\beta} \ \ \textit{and} \ \ Cov[\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})] = \sigma^2\mathbf{P}_{\mathbf{X}_\mathbf{A}}\mathbf{\Sigma}\mathbf{P}_{\mathbf{X}_\mathbf{A}}.$$

(d) *The BLUE of* $\mathbf{X}\boldsymbol{\beta}$ *under* $\mathscr{M}_c$ *can be written as*

$$\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) = (\,\mathbf{I}_n - \mathbf{P}_{\mathbf{X_A}\|\boldsymbol{\Sigma}}\,)\mathbf{X}\mathbf{A}^{+}\mathbf{b} + \mathbf{P}_{\mathbf{X_A}\|\boldsymbol{\Sigma}}\,\mathbf{y} \tag{15}$$

with $E[\,\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})\,] = \mathbf{X}\boldsymbol{\beta}$ *and* $Cov[\,\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})\,] = \sigma^2\mathbf{P}_{\mathbf{X_A}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma}\mathbf{P}'_{\mathbf{X_A}\|\boldsymbol{\Sigma}}$*, where* $\mathbf{P}_{\mathbf{X_A}\|\boldsymbol{\Sigma}}$ *is a solution to the consistent equation* $\mathbf{Z}_1[\,\mathbf{X_A},\, \boldsymbol{\Sigma}\mathbf{E}_{\mathbf{X_A}}\,] = [\,\mathbf{X_A},\, \mathbf{0}\,]$*, and the two block matrices satisfy*

$$\mathscr{R}[\,\mathbf{X_A},\, \boldsymbol{\Sigma}\mathbf{E}_{\mathbf{X_A}}\,] = \mathscr{R}[\,\mathbf{X_A},\, \boldsymbol{\Sigma}\,] \ \ \textit{and} \ \ \mathscr{R}\begin{bmatrix} \mathbf{X}'_{\mathbf{A}} \\ \mathbf{0} \end{bmatrix} \subseteq \mathscr{R}\begin{bmatrix} \mathbf{X}'_{\mathbf{A}} \\ \mathbf{F}_{\mathbf{X_A}}\boldsymbol{\Sigma} \end{bmatrix}. \tag{16}$$

The general expression of $\mathbf{P}_{\mathbf{X_A}\|\boldsymbol{\Sigma}}$ *can be written as*

$$\mathbf{P}_{\mathbf{X_A}\|\boldsymbol{\Sigma}} = [\,\mathbf{X_A},\, \mathbf{0}\,][\,\mathbf{X_A},\, \boldsymbol{\Sigma}\mathbf{F}_{\mathbf{X_A}}\,]^{+} + \mathbf{V}_1(\mathbf{I}_n - [\,\mathbf{X_A},\, \boldsymbol{\Sigma}\mathbf{F}_{\mathbf{X_A}}\,][\,\mathbf{X_A},\, \boldsymbol{\Sigma}\mathbf{F}_{\mathbf{X_A}}\,]^{+}\,), \tag{17}$$

where $\mathbf{V}_1 \in \mathbb{R}^{n\times n}$ *is arbitrary.* $\mathbf{P}_{\mathbf{X_A}\|\boldsymbol{\Sigma}}$ *is unique if and only if* $r\begin{bmatrix} \mathbf{X} & \boldsymbol{\Sigma} \\ \mathbf{A} & \mathbf{0} \end{bmatrix} = r(\mathbf{A}) + n$; $\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *is unique with probability 1 if and only if*

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{b} \end{bmatrix} \in \mathscr{R}\begin{bmatrix} \mathbf{X} & \boldsymbol{\Sigma} \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \tag{18}$$

holds with probability 1.

Proof The results in (a) and (b) are well known, see, e.g., Puntanen and Styan (1989), Puntanen, Styan and Werner (2000). If (2) is consistent, the general solution of (2) is

$$\boldsymbol{\beta} = \mathbf{A}^{+}\mathbf{b} + \mathbf{F_A}\boldsymbol{\gamma}, \tag{19}$$

where the vector $\boldsymbol{\gamma}$ is arbitrary. Hence,

$$\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{A}^{+}\mathbf{b} + \mathrm{OLSE}(\mathbf{X_A}\boldsymbol{\gamma}) \ \ \text{and} \ \ \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{A}^{+}\mathbf{b} + \mathrm{BLUE}(\mathbf{X_A}\boldsymbol{\gamma}). \tag{20}$$

Substituting (19) into the equation in (1) yields the following reparameterized model

$$\mathbf{z} = \mathbf{X_A}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{21}$$

where $\mathbf{z} = \mathbf{y} - \mathbf{X}\mathbf{A}^{+}\mathbf{b}$ and $\mathbf{X_A} = \mathbf{X}\mathbf{F_A}$. The OLSE of $\boldsymbol{\gamma}$ under (21) can be written as

$$\widehat{\boldsymbol{\gamma}} = \mathbf{X}^{+}_{\mathbf{A}}\mathbf{z} + \mathbf{F}_{\mathbf{X_A}}\mathbf{u},$$

where the vector $\mathbf{u}$ is arbitrary. Substituting this result into (19) gives the OLSE of $\boldsymbol{\beta}$ under (4) as follows

$$\mathrm{OLSE}_{\mathscr{M}_c}(\boldsymbol{\beta}) = \mathbf{A}^{+}\mathbf{b} + \mathbf{F_A}\mathbf{X}^{+}_{1}\mathbf{z} + \mathbf{F_A}\mathbf{F}_{\mathbf{X_A}}\mathbf{u}.$$

Hence,

$$\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}\mathbf{A}^{+}\mathbf{b} + \mathbf{P}_{\mathbf{X_A}}\mathbf{z}(\,\mathbf{I}_n - \mathbf{P}_{\mathbf{X_A}}\,)\mathbf{X}\mathbf{A}^{+}\mathbf{b} + \mathbf{P}_{\mathbf{X_A}}\mathbf{y}.$$

It can be derived from (b) that the BLUE of $\mathbf{X_A}\boldsymbol{\gamma}$ under (21) is $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X_A}\boldsymbol{\gamma}) = \mathbf{P}_{\mathbf{X_A}\|\boldsymbol{\Sigma}}\mathbf{z}$. Substituting this result into the second equality in (18) gives the BLUE of $\mathbf{X}\boldsymbol{\beta}$ under (4)

$$\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}\mathbf{A}^{+}\mathbf{b} + \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X_A}\boldsymbol{\gamma}) = \mathbf{X}\mathbf{A}^{+}\mathbf{b} + \mathbf{P}_{\mathbf{X_A}\|\boldsymbol{\Sigma}}\mathbf{z}, \tag{22}$$

as required for (15). It can be seen from (15) and (22) that $\mathrm{BLUE}_c(\mathbf{X}\boldsymbol{\beta})$ is unique with probability 1 if and only if

$$r[\,\mathbf{z},\, \mathbf{X_A},\, \boldsymbol{\Sigma}\,] = r[\,\mathbf{X_A},\, \boldsymbol{\Sigma}\,] \tag{23}$$

holds with probability 1. Applying Lemma 4(b), $\mathbf{A}\mathbf{A}^{+}\mathbf{b} = \mathbf{b}$ and elementary block matrix operations to both sides of (23) gives

$$\begin{aligned} r[\,\mathbf{z},\, \mathbf{X}_{\mathbf{A}},\, \mathbf{\Sigma}\,] &= r[\,\mathbf{y} - \mathbf{X}\mathbf{A}^{+}\mathbf{b},\, \mathbf{X}\mathbf{F}_{\mathbf{A}},\, \mathbf{\Sigma}\,] \\ &= r\begin{bmatrix} \mathbf{y} - \mathbf{X}\mathbf{A}^{+}\mathbf{b} & \mathbf{X} & \mathbf{\Sigma} \\ \mathbf{0} & \mathbf{A} & \mathbf{0} \end{bmatrix} - r(\mathbf{A}) \\ &= r\begin{bmatrix} \mathbf{y} & \mathbf{X} & \mathbf{\Sigma} \\ \mathbf{b} & \mathbf{A} & \mathbf{0} \end{bmatrix} - r(\mathbf{A}), \\ r[\,\mathbf{X}_1,\, \mathbf{\Sigma}\,] &= r[\,\mathbf{X}\mathbf{F}_{\mathbf{A}},\, \mathbf{\Sigma}\,] = r\begin{bmatrix} \mathbf{X} & \mathbf{\Sigma} \\ \mathbf{A} & \mathbf{0} \end{bmatrix} - r(\mathbf{A}). \end{aligned}$$

Thus, (23) is equivalent to $r\begin{bmatrix} \mathbf{y} & \mathbf{X} & \mathbf{\Sigma} \\ \mathbf{b} & \mathbf{A} & \mathbf{0} \end{bmatrix} = r\begin{bmatrix} \mathbf{X} & \mathbf{\Sigma} \\ \mathbf{A} & \mathbf{0} \end{bmatrix}$, i.e., (18) holds. $\square$

For the parametric vector $\mathbf{X}\boldsymbol{\beta}$, the four linear estimators $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$, $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$, $\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ and $\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ of $\mathbf{X}\boldsymbol{\beta}$ under (3) and (4) are not necessarily equal. It is therefore important to reveal relations among these estimators. In Section 2, we investigate the following six equalities among the four estimators for $\mathbf{X}\boldsymbol{\beta}$ under (3) and (4):

(a1) $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$,

(b1) $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$,

(c1) $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$,

(d1) $\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$,

(e1) $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$,

(f1) $\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$.

In Section 3, we investigate the following four equalities for the estimators of an estimable parametric function $\mathbf{K}\boldsymbol{\beta}$ under (3) and (4):

(a2) $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})$,

(b2) $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$,

(c2) $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$,

(d2) $\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$.

In Section 4, we investigate the following three equalities

(a3) $\mathrm{OLSE}_{\mathscr{M}}(\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_1}(\boldsymbol{\beta})$,

(b3) $\mathrm{BLUE}_{\mathscr{M}}(\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_1}(\boldsymbol{\beta})$,

(c3) $\mathrm{OLSE}_{\mathscr{M}_c}(\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_{1c}}(\boldsymbol{\beta})$.

Some useful rank formulas for partitioned matrices due to Masarglia and Styan (1974) are given below. These rank formulas can be used to simplify various matrix equalities involving inverses and Moore-Penrose inverses.

Lemma 4 *Let* $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{m\times k}$ $\mathbf{C} \in \mathbb{R}^{l\times n}$ *and* $\mathbf{D} \in \mathbb{R}^{l\times k}$. *Then:*

(a) $r[\,\mathbf{A},\, \mathbf{B}\,] = r(\mathbf{A}) + r(\mathbf{E}_{\mathbf{A}}\mathbf{B}) = r(\mathbf{B}) + r(\mathbf{E}_{\mathbf{B}}\mathbf{A})$.

(b) $r\begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} = r(\mathbf{A}) + r(\mathbf{CF_A}) = r(\mathbf{C}) + r(\mathbf{AF_C})$.

(c) $r\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} = r(\mathbf{B}) + r(\mathbf{C}) + r(\mathbf{E_B AF_C})$.

(d) $r[\,\mathbf{A},\, \mathbf{B}\,] = r(\mathbf{A}) \Leftrightarrow \mathbf{E_A B} = \mathbf{0} \Leftrightarrow \mathscr{R}(\mathbf{B}) \subseteq \mathscr{R}(\mathbf{A})$.

(e) $r\begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} = r(\mathbf{A}) \Leftrightarrow \mathbf{CF_A} = \mathbf{0} \Leftrightarrow \mathscr{R}(\mathbf{C}') \subseteq \mathscr{R}(\mathbf{A}')$.

(f) *If* $\mathscr{R}(\mathbf{A}) \cap \mathscr{R}(\mathbf{B}) = \{\mathbf{0}\}$, *then* $(\mathbf{E_A B})^+(\mathbf{E_A B}) = \mathbf{B}^+\mathbf{B}$.

(g) *If* $\mathscr{R}(\mathbf{A}') \cap \mathscr{R}(\mathbf{C}') = \{\mathbf{0}\}$, *then* $(\mathbf{CF_A})(\mathbf{CF_A})^+ = \mathbf{CC}^+$.

(h) *if* $\mathscr{R}(\mathbf{B}) \subseteq \mathscr{R}(\mathbf{A})$ *and* $\mathscr{R}(\mathbf{C}') \subseteq \mathscr{R}(\mathbf{A}')$, *then* $\mathbf{CA^-B}$ *is unique*, $\mathbf{CA^-B} = \mathbf{CA^+B}$, *and*

$$r\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = r(\mathbf{A}) + r(\,\mathbf{D} - \mathbf{CA^+B}\,).$$

2 Relations between the four estimators of $\mathbf{X}\boldsymbol{\beta}$

Relations between $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ and $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ under (3), especially the equality $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ have been extensively investigated in the literature, see, e.g., Baltagi (1989), Puntanan and Styan (1989), Puntanan, Styan and Werner (2000), Qian and Schmidt (2003), and Tian and Wiens (2006) among others. A well-known result on the equality $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ is summarized below.

Theorem 1 *Then the following statements are equivalent*:

(a) $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1.*

(b) $\mathbf{P_X \Sigma E_X} = \mathbf{0}$, *i.e.*, $\mathscr{R}(\mathbf{\Sigma X}) \subseteq \mathscr{R}(\mathbf{X})$.

(c) $\mathbf{P_X \Sigma} = \mathbf{\Sigma P_X}$.

Proof Note that both $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ and $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ are unbiased for $\mathbf{X}\boldsymbol{\beta}$. Then by Lemma 2, $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ holds with probability 1 if and only if $(\,\mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}} - \mathbf{P_X}\,)\mathbf{\Sigma} = \mathbf{0}$. From (13), we obtain

$$\mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma} = [\,\mathbf{X},\, \mathbf{0}\,][\,\mathbf{X},\, \mathbf{\Sigma E_X}\,]^+\mathbf{\Sigma}. \tag{24}$$

It is easy to verify that any matrix $[\,\mathbf{A},\, \mathbf{B}\,]$ has a g-inverse $[\,\mathbf{A},\, \mathbf{B}\,]^- = \begin{bmatrix} \mathbf{A}^+ - \mathbf{A}^+\mathbf{B}(\mathbf{E_A B})^+ \\ (\mathbf{E_A E B})^+ \end{bmatrix}$. Applying this to $[\,\mathbf{X},\, \mathbf{\Sigma E_X}\,]$ gives

$$[\,\mathbf{X},\, \mathbf{\Sigma E_X}\,]^- = \begin{bmatrix} \mathbf{X}^+ - \mathbf{X}^+\mathbf{\Sigma E_X}(\mathbf{E_X \Sigma E_X})^+ \\ (\mathbf{E_X \Sigma E_X})^+ \end{bmatrix}.$$

Hence, we obtain by Lemma 4(h) and (24) the following result

$$\mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma} = [\,\mathbf{X},\, \mathbf{0}\,][\,\mathbf{X},\, \mathbf{\Sigma E_X}\,]^+\mathbf{\Sigma} = [\,\mathbf{X},\, \mathbf{0}\,][\,\mathbf{X},\, \mathbf{\Sigma E_X}\,]^-\mathbf{\Sigma} = \mathbf{P_X \Sigma} - \mathbf{P_X \Sigma}(\mathbf{E_X \Sigma E_X})^+\mathbf{\Sigma}. \tag{25}$$

In this case, so that (24) is equivalent to

$$\mathbf{P_X \Sigma}(\mathbf{E_X \Sigma E_X})^+\mathbf{\Sigma} = \mathbf{0}. \tag{26}$$

Simplify (26) by the following equivalence

$$\mathbf{MN}^+ = \mathbf{0} \Leftrightarrow \mathbf{MN}^+\mathbf{N} = \mathbf{0} \Leftrightarrow \mathbf{MN}'\mathbf{N} = \mathbf{0} \Leftrightarrow \mathbf{MN}' = \mathbf{0} \tag{27}$$

leads to the equivalence of (a), (b) and (c). $\square$

Theorem 2 *Let $\mathscr{M}$ and $\mathscr{M}_c$ be as given in* (3) *and* (4)*. Then:*

(a) $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1 if and only if*

$$\mathscr{R}\begin{bmatrix} \mathbf{X}'\boldsymbol{\Sigma} \\ \mathbf{0} \end{bmatrix} \subseteq \mathscr{R}\begin{bmatrix} \mathbf{X}'\mathbf{X} \\ \mathbf{A} \end{bmatrix}. \tag{28}$$

(b) *If* $\mathscr{R}(\mathbf{X}') \cap \mathscr{R}(\mathbf{A}') = \{\mathbf{0}\}$*, then* $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1.*

(c) *Under the condition* $\mathscr{R}(\mathbf{X}) \subseteq \mathscr{R}(\boldsymbol{\Sigma})$, $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1 if and only if* $\mathscr{R}(\mathbf{X}') \cap \mathscr{R}(\mathbf{A}') = \{\mathbf{0}\}$.

(d) *Under the condition* $r(\mathbf{X}) = p$*. Then* $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1 if and only if* $\mathbf{A}\mathbf{X}^{+}\boldsymbol{\Sigma} = \mathbf{0}$.

(e) *Under the conditions* $\mathscr{R}(\mathbf{X}) \subseteq \mathscr{R}(\boldsymbol{\Sigma})$ *and* $r(\mathbf{X}) = p$, $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1 if and only if* $\mathbf{A} = \mathbf{0}$.

Proof Note that both $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ and $\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ are unbiased for $\mathbf{X}\boldsymbol{\beta}$. Then by Lemma 2, $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ holds with probability 1 if and only if

$$\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} = \mathbf{P}_{\mathbf{X}_{\mathbf{A}}}\boldsymbol{\Sigma}. \tag{29}$$

To simplify this equality, we first find the rank of $\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} - \mathbf{P}_{\mathbf{X}_{\mathbf{A}}}\boldsymbol{\Sigma}$. It is easy to see $\mathbf{P}_{\mathbf{X}_{\mathbf{A}}}\mathbf{P}_{\mathbf{X}} = \mathbf{P}_{\mathbf{X}_{\mathbf{A}}}$. Hence, we find by Lemma 4(a) and (b) that

$$\begin{aligned} r(\, \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} - \mathbf{P}_{\mathbf{X}_{\mathbf{A}}}\boldsymbol{\Sigma}\,) &= r(\, \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} - \mathbf{P}_{\mathbf{X}_{\mathbf{A}}}\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}\,) \\ &= r[\, \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma},\, \mathbf{X}_{\mathbf{A}}\,] - r(\mathbf{X}_{\mathbf{A}}) \\ &= r[\, \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma},\, \mathbf{X}\mathbf{F}_{\mathbf{A}}\,] - r(\mathbf{X}\mathbf{F}_{\mathbf{A}}) \\ &= r\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\boldsymbol{\Sigma} \\ \mathbf{A} & \mathbf{0} \end{bmatrix} - r\begin{bmatrix} \mathbf{X}'\mathbf{X} \\ \mathbf{A} \end{bmatrix}. \end{aligned}$$

Hence, $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ holds with probability 1 if and only if

$$r\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\boldsymbol{\Sigma} \\ \mathbf{A} & \mathbf{0} \end{bmatrix} = r\begin{bmatrix} \mathbf{X}'\mathbf{X} \\ \mathbf{A} \end{bmatrix}, \tag{30}$$

which is equivalent to (28) by Lemma 4(d). Results (b)–(e) are also derived from (30). □

Relations between $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ and $\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ were also studied by several authors, see e.g., Baksalary and Kala (1979), Hallum, Lewis and Boullion (1973), Mathew (1983), and Yang, Wen, Cui and Sun (1987), and there are some disputes among these papers. Our result for the equality between $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ and $\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ is given below.

Theorem 3 (a) $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1 if and only if*

$$r\begin{bmatrix} \mathbf{X} & \boldsymbol{\Sigma} \\ \mathbf{A} & \mathbf{0} \end{bmatrix} = r\begin{bmatrix} \mathbf{X} \\ \mathbf{A} \end{bmatrix} + r[\, \mathbf{X},\, \boldsymbol{\Sigma}\,] - r(\mathbf{X}).$$

(b) *Under the condition* $\mathscr{R}(\mathbf{X}) \cap \mathscr{R}(\boldsymbol{\Sigma}) = \{\mathbf{0}\}$, $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1.*

(c) *Under the condition* $\mathscr{R}(\mathbf{X}') \cap \mathscr{R}(\mathbf{A}') = \{\mathbf{0}\}$, $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1.*

(d) *Under the condition* $\mathscr{R}(\mathbf{X}) \subseteq \mathscr{R}(\boldsymbol{\Sigma})$, $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1 if and only if* $\mathscr{R}(\mathbf{X}') \cap \mathscr{R}(\mathbf{A}') = \{\mathbf{0}\}$.

Proof Because both $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ and $\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ are unbiased for $\mathbf{X}\boldsymbol{\beta}$, it follows from Lemma 2 that $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ holds with probability 1 if and only if

$$\mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma} = \mathbf{P}_{\mathbf{X_A}\|\mathbf{\Sigma}}\mathbf{\Sigma}.$$

Note from (13) that the product $\mathbf{P}_{\mathbf{X_A}\|\mathbf{\Sigma}}\mathbf{\Sigma}$ is

$$\mathbf{P}_{\mathbf{X_A}\|\mathbf{\Sigma}}\mathbf{\Sigma} = [\,\mathbf{X_A},\, \mathbf{0}\,][\,\mathbf{X_A},\, \mathbf{\Sigma E_{X_A}}\,]^{+}\mathbf{\Sigma}. \tag{31}$$

From (24), (25), (8), (12), Lemma 4(h) and elementary block matrix operations, we find that

$$\begin{aligned}
& r(\,\mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma} - \mathbf{P}_{\mathbf{X_A}\|\mathbf{\Sigma}}\mathbf{\Sigma}\,) \\
&= r\{\,[\,\mathbf{X},\, \mathbf{0}\,][\,\mathbf{X},\, \mathbf{\Sigma E_X}\,]^{+}\mathbf{\Sigma} - [\,\mathbf{X_A},\, \mathbf{0}\,][\,\mathbf{X_A},\, \mathbf{\Sigma E_{X_A}}\,]^{+}\mathbf{\Sigma}\,\} \\
&= r\begin{bmatrix} -\mathbf{X} & -\mathbf{\Sigma E_X} & \mathbf{0} & \mathbf{0} & \mathbf{\Sigma} \\ \mathbf{0} & \mathbf{0} & \mathbf{X_A} & \mathbf{\Sigma E_{X_A}} & \mathbf{\Sigma} \\ \mathbf{X} & \mathbf{0} & \mathbf{X_A} & \mathbf{0} & \mathbf{0} \end{bmatrix} - r[\,\mathbf{X},\, \mathbf{\Sigma E_X}\,] - r[\,\mathbf{X_A},\, \mathbf{\Sigma E_{X_A}}\,] \\
&= r\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{X_A} & \mathbf{0} & \mathbf{\Sigma} \\ \mathbf{0} & \mathbf{\Sigma E_X} & \mathbf{0} & \mathbf{\Sigma E_{X_A}} & \mathbf{0} \\ \mathbf{X} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} - r[\,\mathbf{X},\, \mathbf{\Sigma}\,] - r[\,\mathbf{X_A},\, \mathbf{\Sigma}\,] \\
&= r[\,\mathbf{X_A},\, \mathbf{\Sigma}\,] + r[\,\mathbf{\Sigma E_X},\, \mathbf{\Sigma E_{X_A}}\,] + r(\mathbf{X}) - r[\,\mathbf{X},\, \mathbf{\Sigma}\,] - r[\,\mathbf{X_A},\, \mathbf{\Sigma}\,] \\
&= r(\mathbf{\Sigma E_{X_A}}) + r(\mathbf{X}) - r[\,\mathbf{X},\, \mathbf{\Sigma}\,] \\
&= r[\,\mathbf{X_A}\, \mathbf{\Sigma}\,] - r(\mathbf{X_A}) + r(\mathbf{X}) - r[\,\mathbf{X},\, \mathbf{\Sigma}\,] \\
&= r\begin{bmatrix} \mathbf{X} & \mathbf{\Sigma} \\ \mathbf{A} & \mathbf{0} \end{bmatrix} - r\begin{bmatrix} \mathbf{X} \\ \mathbf{A} \end{bmatrix} - r[\,\mathbf{X},\, \mathbf{\Sigma}\,] + r(\mathbf{X}).
\end{aligned}$$

Letting this equality be zero gives (a). Results (b), (c) and (d) follow from (a). □

Mathew (1983) showed that $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ holds with probability 1 if and only if $\mathscr{R}(\mathbf{\Sigma}) \cap \mathscr{R}(\mathbf{X}) \subseteq \mathscr{R}[\,\mathbf{X}(\mathbf{I}_n - \mathbf{A}^{+}\mathbf{A})\,]$. It can be shown by Lemma 4(a) and (b) that this result is equivalent to Theorem 3(a).

Theorem 4 *Let $\mathscr{M}$ and $\mathscr{M}_c$ be as given in* (3) *and* (4). *Then*:

(a) $\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1 if and only if* $(\mathbf{XF_A})(\mathbf{XF_A})^{+}\mathbf{\Sigma} = \mathbf{\Sigma}(\mathbf{XF_A})(\mathbf{XF_A})^{+}$, *or equivalently*,

$$r\begin{bmatrix} \mathbf{\Sigma X} & \mathbf{X} \\ \mathbf{0} & \mathbf{A} \\ \mathbf{A} & \mathbf{0} \end{bmatrix} = r\begin{bmatrix} \mathbf{X} \\ \mathbf{A} \end{bmatrix} + r(\mathbf{A}).$$

(b) *Under the condition* $\mathscr{R}(\mathbf{X}') \cap \mathscr{R}(\mathbf{A}') = \{\mathbf{0}\}$, $\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1 f and only if* $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$.

Proof The proof is left for the reader. □

Theorem 5 *Let $\mathscr{M}$ and $\mathscr{M}_c$ be as given in* (3) *and* (4). *Then*:

(a) $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1 if and only if* $\mathscr{R}\begin{bmatrix} \mathbf{\Sigma X} \\ \mathbf{0} \end{bmatrix} \subseteq \mathscr{R}\begin{bmatrix} \mathbf{X} \\ \mathbf{A} \end{bmatrix}$. *In this case*, $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1.*

(b) *Under the condition* $\mathscr{R}(\mathbf{X}') \cap \mathscr{R}(\mathbf{A}') = \{\mathbf{0}\}$, $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds if and only if* $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1. In this case*,

$$\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$$

holds with probability 1.

Proof Since both $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ and $\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ are unbiased for $\mathbf{X}\boldsymbol{\beta}$, it follows from Lemma 2 that $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ holds with probability 1 if and only if $\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} = \mathbf{P}_{\mathbf{X}_{\mathbf{A}}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma}$. From (31), Lemma 4(a), (b) and (h), (15) and elementary block matrix operations, we find that

$$\begin{aligned} r(\,\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} - \mathbf{P}_{\mathbf{X}_{\mathbf{A}}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma}\,) &= r(\,\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} - [\,\mathbf{X}_{\mathbf{A}},\, \mathbf{0}\,][\,\mathbf{X}_{\mathbf{A}},\, \boldsymbol{\Sigma}\mathbf{E}_{\mathbf{X}_{\mathbf{A}}}\,]^{+}\boldsymbol{\Sigma}\,) \\ &= r\begin{bmatrix} \mathbf{X}_{\mathbf{A}} & \boldsymbol{\Sigma}\mathbf{E}_{\mathbf{X}_{\mathbf{A}}} & \boldsymbol{\Sigma} \\ \mathbf{X}_{\mathbf{A}} & \mathbf{0} & \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} \end{bmatrix} - r[\,\mathbf{X}_{\mathbf{A}},\, \boldsymbol{\Sigma}\mathbf{E}_{\mathbf{X}_{\mathbf{A}}}\,] \\ &= r\begin{bmatrix} \mathbf{X}_{\mathbf{A}} & \mathbf{0} & \boldsymbol{\Sigma} \\ \mathbf{0} & -\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}\mathbf{E}_{\mathbf{X}_{\mathbf{A}}} & \mathbf{0} \end{bmatrix} - r[\,\mathbf{X}_{\mathbf{A}},\, \boldsymbol{\Sigma}\,] \\ &= r(\mathbf{E}_{\mathbf{X}_{\mathbf{A}}}\boldsymbol{\Sigma}\mathbf{P}_{\mathbf{X}}) \\ &= r[\,\mathbf{X}_{\mathbf{A}},\, \boldsymbol{\Sigma}\mathbf{X}\,] - r(\mathbf{X}_{\mathbf{A}}) \\ &= r[\,\mathbf{X}\mathbf{F}_{\mathbf{A}},\, \boldsymbol{\Sigma}\mathbf{X}\,] - r(\mathbf{X}\mathbf{F}_{\mathbf{A}}) \\ &= r\begin{bmatrix} \mathbf{X} & \boldsymbol{\Sigma}\mathbf{X} \\ \mathbf{A} & \mathbf{0} \end{bmatrix} - r\begin{bmatrix} \mathbf{X} \\ \mathbf{A} \end{bmatrix}. \end{aligned}$$

Letting this equality be zero yields the desired results. □

Theorem 6 *Let $\mathscr{M}$ and $\mathscr{M}_c$ be as given in* (3) *and* (4). *Then* $\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1 if and only if*

$$r\begin{bmatrix} \mathbf{X} & \boldsymbol{\Sigma} \\ \mathbf{A} & \mathbf{0} \end{bmatrix} = r\begin{bmatrix} \mathbf{X} \\ \mathbf{A} \end{bmatrix} + r[\,\mathbf{X},\, \boldsymbol{\Sigma}\,] - r(\mathbf{X}) \quad \textit{and} \quad r\begin{bmatrix} \boldsymbol{\Sigma}\mathbf{X} & \mathbf{X} \\ \mathbf{A} & \mathbf{0} \end{bmatrix} = r(\mathbf{X}) + r(\mathbf{A}).$$

In this case, $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta})$ *holds with probability 1.*

Proof The proof is left for the reader. □

3 Relations between estimators of parametric functions

Let $\mathbf{K}$ be an $l \times p$ known matrix. Then the product $\mathbf{K}\boldsymbol{\beta}$ is called a parametric function of $\boldsymbol{\beta}$ under (3). When $\mathbf{K} = \mathbf{I}_p$, $\mathbf{K}\boldsymbol{\beta} = \boldsymbol{\beta}$ is the unknown parameter vector under (3). The parametric function $\mathbf{K}\boldsymbol{\beta}$ is said to be estimable under (3) if there exists an $l \times n$ matrix $\mathbf{L}$ so that $E(\mathbf{L}\mathbf{y}) = \mathbf{K}\boldsymbol{\beta}$ holds. It is well known that $\mathbf{K}\boldsymbol{\beta}$ is estimable under (3) if and only if $\mathscr{R}(\mathbf{K}') \subseteq \mathscr{R}(\mathbf{X}')$, see Alalouf and Styan (1979).

Since there is a linear restriction for $\boldsymbol{\beta}$ under (4), the estimability of $\mathbf{K}\boldsymbol{\beta}$ under (4) is defined be below.

Definition 3 Let $\mathbf{K}$ be an $l \times p$ known matrix. The vector $\mathbf{K}\boldsymbol{\beta}$ is said to be estimable under $\mathscr{M}_c$ in (4) if there exist an $l \times n$ constant matrix $\mathbf{L}$ and an $l \times 1$ constant vector $\mathbf{c}$ such that $E(\mathbf{L}\mathbf{y} + \mathbf{c}) = \mathbf{K}\boldsymbol{\beta}$.

Note $E(\mathbf{L}\mathbf{y}+\mathbf{c}) = \mathbf{L}\mathbf{X}\boldsymbol{\beta}+\mathbf{c}$. Hence $E(\mathbf{L}\mathbf{y}+\mathbf{c}) = \mathbf{K}\boldsymbol{\beta}$ is equivalent to $\mathbf{L}\mathbf{X}\boldsymbol{\beta}+\mathbf{c} = \mathbf{K}\boldsymbol{\beta}$. However, we cannot simply say that $\mathbf{L}\mathbf{X} = \mathbf{K}$ and $\mathbf{c} = \mathbf{0}$, because $\boldsymbol{\beta}$ is not free but subject to the linear equation $\mathbf{A}\boldsymbol{\beta} = \mathbf{b}$. A well-known result on the estimability of $\mathbf{K}\boldsymbol{\beta}$ under (4) asserts that

$$\mathbf{K}\boldsymbol{\beta} \text{ is estimable under (4) if and only if } \mathscr{R}(\mathbf{K}') \subseteq \mathscr{R}[\,\mathbf{X}',\, \mathbf{A}'\,]. \tag{32}$$

It can be seen from (32) that if $\mathbf{K}\boldsymbol{\beta}$ is estimable under (3), then $\mathbf{K}\boldsymbol{\beta}$ is estimable under (4).

Let $\mathscr{M}$ be as given in (3). If $\mathbf{K} = \mathbf{K}_1\mathbf{K}_2$, then

$$\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{K}_1\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}_2\boldsymbol{\beta}), \tag{33}$$

see Groß, Trenkler and Werner (2001). Suppose $\mathbf{K}\boldsymbol{\beta}$ is estimable under (3), i.e., $\mathbf{K} = \mathbf{L}_1\mathbf{X}$ for some $\mathbf{L}_1$. Then the BLUE of $\mathbf{K}\boldsymbol{\beta}$ is a linear estimator $\mathbf{Ly}$ such that $E(\mathbf{Ly}) = \mathbf{K}\boldsymbol{\beta}$ and for any other linear unbiased estimator $\mathbf{My}$ of $\mathbf{K}\boldsymbol{\beta}$, $Cov(\mathbf{My}) - Cov(\mathbf{Ly}) = \mathbf{M}\boldsymbol{\Sigma}\mathbf{M}' - \mathbf{L}\boldsymbol{\Sigma}\mathbf{L}'$ is nonnegative. It is easy to verify that if $\mathbf{K}\boldsymbol{\beta}$ is estimable under (3), then

$$\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{L}\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}), \tag{34}$$

see, e.g., Groß, Trenkler and Werner (2001).

Results (33) and (34) imply that the OLSEs and BLUEs of the estimable parametric function $\mathbf{K}\boldsymbol{\beta}$ under (3) and (4) can be derived from the OLSEs and BLUEs of $\mathbf{X}\boldsymbol{\beta}$ under (3) and (4).

Suppose $\mathbf{K}\boldsymbol{\beta}$ is estimable under (4), i.e., there exists a $\mathbf{K} = [\,\mathbf{L}_1,\ \mathbf{L}_2\,]$ such that $\mathbf{L} = \mathbf{L}_1\mathbf{X} + \mathbf{L}_2\mathbf{A}$. From (33) and (34), we see

$$\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{L}_1\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) + \mathbf{L}_2\mathbf{b}, \quad \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{L}_1\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{X}\boldsymbol{\beta}) + \mathbf{L}_2\mathbf{b}. \tag{35}$$

From these expressions and Lemma 3, we obtain the following two consequences.

Theorem 7 *Let $\mathscr{M}$ be as given in* (3) *and suppose $\mathbf{K}\boldsymbol{\beta}$ is estimable under $\mathscr{M}$, i.e., $\mathbf{K}$ can be written as $\mathbf{K} = \mathbf{LX}$. Then*:

(a) $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{LP_X y} = \mathbf{KX^+y}$ *with* $E[\,\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})\,] = \mathbf{K}\boldsymbol{\beta}$ *and* $Cov[\,\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})\,] = \sigma^2\mathbf{KX}^+\boldsymbol{\Sigma}(\mathbf{KX}^+)' = \sigma^2(\mathbf{LP_X})\boldsymbol{\Sigma}(\mathbf{LP_X})'$.

(b) $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{Sy}$, *where $\mathbf{S}$ is the solution of the consistent equation* $\mathbf{S}[\,\mathbf{X},\ \mathbf{VE_X}\,] = [\,\mathbf{K},\ \mathbf{0}\,]$. *Moreover,* $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})$ *can be written as*

$$\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{L}\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{LP}_{\mathbf{X}\|\boldsymbol{\Sigma}}\mathbf{y}$$

with $E[\,\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})\,] = \mathbf{K}\boldsymbol{\beta}$ *and* $Cov[\,\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})\,] = \sigma^2(\mathbf{LP}_{\mathbf{X}\|\boldsymbol{\Sigma}})\boldsymbol{\Sigma}(\mathbf{LP}_{\mathbf{X}\|\boldsymbol{\Sigma}})'$.

Theorem 8 *Let $\mathscr{M}_c$ be as given in* (4) *and suppose $\mathbf{K}\boldsymbol{\beta}$ is estimable under* (4), *i.e., $\mathbf{K}$ can be written as $\mathbf{K} = \mathbf{L}_1\mathbf{X} + \mathbf{L}_2\mathbf{A}$.*

(a) *The OLSE of $\mathbf{K}\boldsymbol{\beta}$ under $\mathscr{M}_c$ can be written as*

$$\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{L}_1(\,\mathbf{I}_n - \mathbf{P}_{\mathbf{X_A}}\,)\mathbf{XA}^+\mathbf{b} + \mathbf{L}_1\mathbf{P}_{\mathbf{X_A}}\mathbf{y} + \mathbf{L}_2\mathbf{b},$$

where $\mathbf{X_A} = \mathbf{XF_A}$. In this case, $E[\,\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})\,] = \mathbf{K}\boldsymbol{\beta}$ *and* $Cov[\,\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})\,] = \sigma^2\mathbf{L}_1\mathbf{P}_{\mathbf{X_A}}\boldsymbol{\Sigma}\mathbf{P}_{\mathbf{X_A}}\mathbf{L}_1'$.

(b) *The BLUE of $\mathbf{K}\boldsymbol{\beta}$ under $\mathscr{M}_c$ can be written as*

$$\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{L}_1(\,\mathbf{I}_n - \mathbf{P}_{\mathbf{X_A}\|\boldsymbol{\Sigma}}\,)\mathbf{XA}^+\mathbf{b} + \mathbf{L}_1\mathbf{P}_{\mathbf{X_A}\|\boldsymbol{\Sigma}}\mathbf{y} + \mathbf{L}_2\mathbf{b}$$

with $E[\,\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})\,] = \mathbf{K}\boldsymbol{\beta}$ *and* $Cov[\,\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})\,] = \sigma^2(\mathbf{L}_1\mathbf{P}_{\mathbf{X_A}\|\boldsymbol{\Sigma}})\boldsymbol{\Sigma}(\mathbf{L}_1\mathbf{P}_{\mathbf{X_A}\|\boldsymbol{\Sigma}})'$.

Concerning the relations among the four estimators for $\mathbf{K}\boldsymbol{\beta}$ in Theorems 7 and 8. We have the following several results.

Theorem 9 *Let $\mathscr{M}$ be as given in* (3), *and suppose $\mathbf{K}\boldsymbol{\beta}$ is estimable under $\mathscr{M}$, i.e., $\mathbf{K}$ can be written as $\mathbf{K} = \mathbf{LX}$. Then* $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})$ *holds with probability 1 if and only if* $\mathscr{R}[(\mathbf{KX}^+\boldsymbol{\Sigma})'] \subseteq \mathscr{R}(\mathbf{X})$.

Proof Since both $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})$ and $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})$ are unbiased for $\mathbf{K}\boldsymbol{\beta}$ under $\mathscr{M}$, it follows from Lemma 2, Theorem 7(a) and (b) that $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})$ holds with probability 1 if and only if $\mathbf{LP_X}\boldsymbol{\Sigma} = \mathbf{LP}_{\mathbf{X}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma}$. From (25), we see

$$\mathbf{LP_X}\boldsymbol{\Sigma} - \mathbf{LP}_{\mathbf{X}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma} = \mathbf{LP_X}\boldsymbol{\Sigma}(\mathbf{E_X}\boldsymbol{\Sigma}\mathbf{E_X})^+\boldsymbol{\Sigma}.$$

Hence, $\mathbf{LP_X}\boldsymbol{\Sigma} = \mathbf{LP}_{\mathbf{X}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma}$ is equivalent to $\mathbf{LP_X}\boldsymbol{\Sigma}(\mathbf{E_X}\boldsymbol{\Sigma}\mathbf{E_X})^+\boldsymbol{\Sigma} = \mathbf{0}$. It is easy to verify by (27) that this equality is equivalent to $\mathbf{LP_X}\boldsymbol{\Sigma}\mathbf{E_X} = \mathbf{0}$, which can also be written as $\mathbf{KX}^+\boldsymbol{\Sigma} = \mathbf{KX}^+\boldsymbol{\Sigma}\mathbf{P_X}$. □

Theorem 10 *Let $\mathscr{M}$ and $\mathscr{M}_c$ be as given in* (3) *and* (4), *and suppose* $\mathbf{K}\boldsymbol{\beta}$ *is estimable under* $\mathscr{M}$, *i.e.*, $\mathbf{K} = \mathbf{L}\mathbf{X}$. *Then*:

(a) $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$ *holds with probability 1 if and only if* $\mathbf{K}\mathbf{X}^{+}\boldsymbol{\Sigma} = \mathbf{K}(\mathbf{X}\mathbf{F}_{\mathbf{A}})^{+}\boldsymbol{\Sigma}$.

(b) *Under the condition* $\mathscr{R}(\mathbf{X}') \cap \mathscr{R}(\mathbf{A}') = \{\mathbf{0}\}$, $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$ *holds with probability 1.*

(c) *Under the condition* $\mathscr{R}(\mathbf{X}) \subseteq \mathscr{R}(\boldsymbol{\Sigma})$, $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$ *holds with probability 1 if and only if* $\mathscr{R}\begin{bmatrix} \mathbf{K}' \\ \mathbf{0} \end{bmatrix} \subseteq \mathscr{R}\begin{bmatrix} \mathbf{X}'\mathbf{X} \\ \mathbf{A} \end{bmatrix}$.

(d) *Under the conditions* $\mathscr{R}(\mathbf{X}) \subseteq \mathscr{R}(\boldsymbol{\Sigma})$ *and* $r(\mathbf{X}) = p$, $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$ *holds with probability 1 if and only if* $\mathbf{A}(\mathbf{X}'\mathbf{X})^{+}\mathbf{K}' = \mathbf{0}$.

Proof Since both $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})$ and $\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$ are unbiased for $\mathbf{K}\boldsymbol{\beta}$, it follows from Lemma 2, Theorems 7(a) and 8(a) that $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$ if and only if

$$\mathbf{L}\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} = \mathbf{L}\mathbf{P}_{\mathbf{X}_{\mathbf{A}}}\boldsymbol{\Sigma},$$

that is, $\mathbf{K}\mathbf{X}^{+}\boldsymbol{\Sigma} = \mathbf{K}\mathbf{E}_{\mathbf{A}}(\mathbf{X}\mathbf{E}_{\mathbf{A}})^{+}\boldsymbol{\Sigma}$. If $\mathscr{R}(\mathbf{X}')\cap\mathscr{R}(\mathbf{A}') = \{\mathbf{0}\}$, then $(\mathbf{X}\mathbf{E}_{\mathbf{A}})(\mathbf{X}\mathbf{E}_{\mathbf{A}})^{+} = \mathbf{X}\mathbf{X}^{+}$. Hence, (b) follows. The proofs of (c) and (d) are left for the reader. □

Theorem 11 *Let $\mathscr{M}$ and $\mathscr{M}_c$ be as given in* (3) *and* (4), *and suppose* $\mathbf{K}\boldsymbol{\beta}$ *is estimable under* $\mathscr{M}$, *i.e.*, $\mathbf{K} = \mathbf{L}\mathbf{X}$. *Then*:

(a) $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$ *holds with probability 1 if and only if*

$$r\begin{bmatrix} \mathbf{0} & \mathbf{X}' & \mathbf{K}' \\ \mathbf{X} & \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{A} & \mathbf{0} & \mathbf{0} \end{bmatrix} = r\begin{bmatrix} \mathbf{X} \\ \mathbf{A} \end{bmatrix} + r[\,\mathbf{X},\, \boldsymbol{\Sigma}\,].$$

(b) *Under the condition* $\mathscr{R}(\mathbf{X}) \subseteq \mathscr{R}(\boldsymbol{\Sigma})$, $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$ *holds with probability 1 if and only if*

$$\mathscr{R}\begin{bmatrix} \mathbf{K}' \\ \mathbf{0} \end{bmatrix} \subseteq \mathscr{R}\begin{bmatrix} \mathbf{X}'\boldsymbol{\Sigma}^{+}\mathbf{X} \\ \mathbf{A} \end{bmatrix}.$$

(c) *Under the conditions that* $\boldsymbol{\Sigma}$ *is positive definite and* $r(\mathbf{X}) = p$, $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$ *holds with probability 1 if and only if* $\mathbf{A}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{K} = \mathbf{0}$.

Proof It can be seen from Lemma 2, Theorems 7(b) and 8(b) that $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$ holds with probability 1 if and only if $\mathbf{L}\mathbf{P}_{\mathbf{X}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma} = \mathbf{L}\mathbf{P}_{\mathbf{X}_{\mathbf{A}}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma}$, this is,

$$\mathbf{L}[\,\mathbf{X},\, \mathbf{0}\,][\,\mathbf{X},\, \boldsymbol{\Sigma}\mathbf{E}_{\mathbf{X}}\,]^{+}\boldsymbol{\Sigma} - \mathbf{L}[\,\mathbf{X}_{\mathbf{A}},\, \mathbf{0}\,][\,\mathbf{X}_{\mathbf{A}},\, \boldsymbol{\Sigma}\mathbf{E}_{\mathbf{X}_{\mathbf{A}}}\,]^{+}\boldsymbol{\Sigma} = \mathbf{0}.$$

Applying Lemma 4(h) to the left-hand side gives

$$r(\,\mathbf{L}[\,\mathbf{X},\, \mathbf{0}\,][\,\mathbf{X},\, \boldsymbol{\Sigma}\mathbf{E}_{\mathbf{X}}\,]^{+}\boldsymbol{\Sigma} - \mathbf{L}[\,\mathbf{X}_{\mathbf{A}},\, \mathbf{0}\,][\,\mathbf{X}_{\mathbf{A}},\, \boldsymbol{\Sigma}\mathbf{E}_{\mathbf{X}_{\mathbf{A}}}\,]^{+}\boldsymbol{\Sigma}\,) = r\begin{bmatrix} \mathbf{0} & \mathbf{X}' & \mathbf{K}' \\ \mathbf{X} & \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{A} & \mathbf{0} & \mathbf{0} \end{bmatrix} - r\begin{bmatrix} \mathbf{X} \\ \mathbf{A} \end{bmatrix} - r[\,\mathbf{X},\, \boldsymbol{\Sigma}\,].$$

The verification of the rank equality is left for the reader. □

Theorem 12 *Let $\mathscr{M}_c$ be as given in* (4), *and suppose* $\mathbf{K}\boldsymbol{\beta}$ *is estimable under* $\mathscr{M}_c$, *i.e.*, $\mathbf{K} = \mathbf{L}_1\mathbf{X} + \mathbf{L}_2\mathbf{A}$. *Then* $\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$ *holds with probability 1 if and only if* $\mathbf{K}\mathbf{X}_{\mathbf{A}}^{+}\boldsymbol{\Sigma} = \mathbf{K}\mathbf{X}_{\mathbf{A}}^{+}\boldsymbol{\Sigma}\mathbf{P}_{\mathbf{X}_{\mathbf{A}}}$.

Proof Since both $\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$ and $\mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$ are unbiased for $\mathbf{K}\boldsymbol{\beta}$, it follows from Lemma 2, Theorem 8(a) and (b) that $\mathrm{OLSE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_c}(\mathbf{K}\boldsymbol{\beta})$ holds with probability 1 if and only if $\mathbf{L}\mathbf{P}_{\mathbf{X}_\mathbf{A}}\boldsymbol{\Sigma} = \mathbf{L}\mathbf{P}_{\mathbf{X}_\mathbf{A}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma}$, i.e.,

$$\mathbf{L}\mathbf{P}_{\mathbf{X}_\mathbf{A}}\boldsymbol{\Sigma} - \mathbf{L}[\,\mathbf{X}_\mathbf{A},\,\mathbf{0}\,][\,\mathbf{X}_\mathbf{A},\,\boldsymbol{\Sigma}\mathbf{E}_{\mathbf{X}_\mathbf{A}}\,]^{+}\boldsymbol{\Sigma} = \mathbf{0}.$$

From (25)

$$\mathbf{L}\mathbf{P}_{\mathbf{X}_\mathbf{A}}\boldsymbol{\Sigma} - \mathbf{L}\mathbf{P}_{\mathbf{X}_\mathbf{A}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma} = \mathbf{L}\mathbf{P}_{\mathbf{X}_\mathbf{A}}\boldsymbol{\Sigma}(\mathbf{E}_{\mathbf{X}_\mathbf{A}}\boldsymbol{\Sigma}\mathbf{E}_{\mathbf{X}_\mathbf{A}})^{+}\boldsymbol{\Sigma}.$$

Hence, $\mathbf{L}\mathbf{P}_{\mathbf{X}_\mathbf{A}}\boldsymbol{\Sigma} = \mathbf{L}\mathbf{P}_{\mathbf{X}_\mathbf{A},\boldsymbol{\Sigma}}\boldsymbol{\Sigma}$ is equivalent to $\mathbf{L}\mathbf{P}_{\mathbf{X}_\mathbf{A}}\boldsymbol{\Sigma}(\mathbf{E}_{\mathbf{X}_\mathbf{A}}\boldsymbol{\Sigma}\mathbf{E}_{\mathbf{X}_\mathbf{A}})^{+}\boldsymbol{\Sigma} = \mathbf{0}$. It is easy to verify by (27) that this equality is equivalent to $\mathbf{L}\mathbf{P}_{\mathbf{X}_\mathbf{A}}\boldsymbol{\Sigma}\mathbf{E}_{\mathbf{X}_\mathbf{A}} = \mathbf{0}$, which can also be written as $\mathbf{K}\mathbf{X}_\mathbf{A}^{+}\boldsymbol{\Sigma} = \mathbf{K}\mathbf{X}_\mathbf{A}^{+}\boldsymbol{\Sigma}\mathbf{P}_{\mathbf{X}_\mathbf{A}}$. $\square$

4 Relations between estimators for the nonsingular linear model and its sub-sample model

We have seen that the model (7) is the first part of the model (3). In many situations, a general linear model is the combination of a set of sub-sample models, for example, the fix-effect model for panel data in econometrics, see, Gourieroux and Monfort (1980), Qian and Schmidt (2003).

If $r(\mathbf{X}) = r(\mathbf{X}_{(1)}) = p$ and $r(\boldsymbol{\Sigma}) = n$, (3), (4), (7) and (8) are said to be nonsingular models. In such cases, $\mathbf{X}^{+}\mathbf{X} = \mathbf{I}_p$, and $\boldsymbol{\beta}$ is estimable under (3), (4), (7) and (8). From Lemma 3, we obtain the OLSEs and BLUEs of $\boldsymbol{\beta}$ under (3), (4), (7) and (8) as follows.

Lemma 5 *Let $\mathscr{M}$ and $\mathscr{M}_c$ be as given in* (3) *and* (4), *and suppose $r(\mathbf{X}) = p$ and $r(\boldsymbol{\Sigma}) = n$. Then*:

(a) *The OLSE of $\boldsymbol{\beta}$ under $\mathscr{M}$ can be written as $\mathrm{OLSE}_{\mathscr{M}}(\boldsymbol{\beta}) = \mathbf{X}^{+}\mathbf{y}$ with $E[\,\mathrm{OLSE}_{\mathscr{M}}(\boldsymbol{\beta})\,] = \boldsymbol{\beta}$ and $Cov[\,\mathrm{OLSE}_{\mathscr{M}}(\boldsymbol{\beta})\,] = \sigma^2\mathbf{X}^{+}\boldsymbol{\Sigma}(\mathbf{X}^{+})'$.*

(b) *The BLUE of $\boldsymbol{\beta}$ under $\mathscr{M}$ can be written as $\mathrm{BLUE}_{\mathscr{M}}(\boldsymbol{\beta}) = (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{y}$ with $E[\,\mathrm{BLUE}_{\mathscr{M}}(\boldsymbol{\beta})\,] = \boldsymbol{\beta}$ and $Cov[\,\mathrm{BLUE}_{\mathscr{M}}(\boldsymbol{\beta})\,] = \sigma^2(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$.*

(c) *The OLSE of $\boldsymbol{\beta}$ under $\mathscr{M}_c$ can be written as*

$$\mathrm{OLSE}_{\mathscr{M}_c}(\boldsymbol{\beta}) = (\,\mathbf{I}_p - \mathbf{X}_\mathbf{A}^{+}\mathbf{X}\,)\mathbf{A}^{+}\mathbf{b} + \mathbf{X}_\mathbf{A}^{+}\mathbf{y},$$

where $\mathbf{X}_\mathbf{A} = \mathbf{X}\mathbf{F}_\mathbf{A}$. In this case, $E[\,\mathrm{OLSE}_{\mathscr{M}_c}(\boldsymbol{\beta})\,] = \boldsymbol{\beta}$ and $Cov[\,\mathrm{OLSE}_{\mathscr{M}_c}(\boldsymbol{\beta})\,] = \sigma^2\mathbf{X}_\mathbf{A}^{+}\boldsymbol{\Sigma}(\mathbf{X}_\mathbf{A}^{+})'$.

Lemma 6 *Let $\mathscr{M}_1$ and $\mathscr{M}_{1c}$ be as given in* (7) *and* (8), *and suppose $r(\mathbf{X}_{(1)}) = p$ and $r(\boldsymbol{\Sigma}) = n$. Then*:

(a) *The OLSE of $\boldsymbol{\beta}$ under $\mathscr{M}_1$ can be written as $\mathrm{OLSE}_{\mathscr{M}_1}(\boldsymbol{\beta}) = \mathbf{X}_{(1)}^{+}\mathbf{y}_1$ with $E[\,\mathrm{OLSE}_{\mathscr{M}_1}(\boldsymbol{\beta})\,] = \boldsymbol{\beta}$ and $Cov[\,\mathrm{OLSE}_{\mathscr{M}_1}(\boldsymbol{\beta})\,] = \sigma^2\mathbf{X}_{(1)}^{+}\boldsymbol{\Sigma}_{11}(\mathbf{X}_{(1)}^{+})'$.*

(b) *The BLUE of $\boldsymbol{\beta}$ under $\mathscr{M}_1$ can be written as $\mathrm{BLUE}_{\mathscr{M}_1}(\boldsymbol{\beta}) = (\mathbf{X}_{(1)}'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{X}_{(1)})^{-1}\mathbf{X}_{(1)}'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{y}_1$ with $E[\,\mathrm{BLUE}_{\mathscr{M}_1}(\boldsymbol{\beta})\,] = \boldsymbol{\beta}$ and $Cov[\,\mathrm{BLUE}_{\mathscr{M}_1}(\boldsymbol{\beta})\,] = \sigma^2(\mathbf{X}_{(1)}'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{X}_{(1)})^{-1}$.*

(c) *The OLSE of $\boldsymbol{\beta}$ under $\mathscr{M}_{1c}$ can be written as*

$$\mathrm{OLSE}_{\mathscr{M}_{1c}}(\boldsymbol{\beta}) = [\,\mathbf{I}_{p_1} - (\mathbf{X}_{(1)})_\mathbf{A}^{+}\mathbf{X}_{(1)}\,]\mathbf{A}^{+}\mathbf{b} + (\mathbf{X}_{(1)})_\mathbf{A}^{+}\mathbf{y}_1,$$

where $(\mathbf{X}_{(1)})_\mathbf{A} = \mathbf{X}_{(1)}\mathbf{F}_\mathbf{A}$. In this case, $E[\,\mathrm{OLSE}_{\mathscr{M}_{1c}}(\boldsymbol{\beta})\,] = \boldsymbol{\beta}$ and $Cov[\,\mathrm{OLSE}_{\mathscr{M}_{1c}}(\boldsymbol{\beta})\,] = \sigma^2(\mathbf{X}_{(1)})_\mathbf{A}^{+}\boldsymbol{\Sigma}_{11}[(\mathbf{X}_{(1)})_\mathbf{A}^{+}]'$.

The following results are derived from Lemmas 6 and 7.

Theorem 13 *Let $\mathscr{M}$ and $\mathscr{M}_c$ be as given in* (3) *and* (4), *and suppose* $r(\mathbf{X}_{(1)}) = p$ *and* $r(\boldsymbol{\Sigma}) = n$. *Then* $\mathrm{OLSE}_{\mathscr{M}}(\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_1}(\boldsymbol{\beta})$ *holds with probability 1 if and only if* $\mathbf{X}_{(2)} = \mathbf{0}$.

Proof Note that both $\mathrm{BLUE}_{\mathscr{M}}(\boldsymbol{\beta})$ and $\mathrm{OLSE}_{\mathscr{M}_1}(\boldsymbol{\beta})$ are unbiased for $\boldsymbol{\beta}$. Also note that $\mathbf{y}_1 = [\,\mathbf{I}_{n_1},\ \mathbf{0}\,]\mathbf{y}$ and $\boldsymbol{\Sigma}$ is nonsingular. Hence we find from Lemmas 4, 6(a) and 7(a) that $\mathrm{OLSE}_{\mathscr{M}}(\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_1}(\boldsymbol{\beta})$ holds with probability 1 if and only if $\mathbf{X}^{+} = [\,\mathbf{X}^{+}_{(1)}, \mathbf{0}\,]$, which is equivalent to $\mathbf{X} = \begin{bmatrix} \mathbf{X}_{(1)} \\ \mathbf{0} \end{bmatrix}$. $\square$

Theorem 14 *Let $\mathscr{M}$ and $\mathscr{M}_c$ be as given in* (3) *and* (4), *and suppose* $r(\mathbf{X}_{(1)}) = p$ *and* $r(\boldsymbol{\Sigma}) = n$. *Then the following statements are equivalent*:

(a) $\mathrm{BLUE}_{\mathscr{M}}(\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_1}(\boldsymbol{\beta})$ *holds with probability 1.*

(b) $\mathbf{X}_{(2)} = \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{X}_{(1)}$.

(c) $r\begin{bmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{X}_{(1)} \\ \boldsymbol{\Sigma}_{21} & \mathbf{X}_{(2)} \end{bmatrix} = r(\boldsymbol{\Sigma}_{11})$.

Proof Note that both $\mathrm{OLSE}_{\mathscr{M}}(\boldsymbol{\beta})$ and $\mathrm{BLUE}_{\mathscr{M}_1}(\boldsymbol{\beta})$ are unbiased for $\boldsymbol{\beta}$. Then we see from Lemma 2, Lemmas 6(b) and 7(b) that $\mathrm{BLUE}_{\mathscr{M}}(\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}_1}(\boldsymbol{\beta})$ holds with probability 1 if and only if

$$(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}' = [\,(\mathbf{X}'_{(1)}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{X}_{(1)})^{-1}\mathbf{X}'_{(1)}\boldsymbol{\Sigma}_{11}^{-1},\, \mathbf{0}\,]\boldsymbol{\Sigma}. \tag{36}$$

This is equivalent to

$$\mathbf{X}' = \mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X}(\mathbf{X}'_{(1)}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{X}_{(1)})^{-1}[\,\mathbf{X}'_{(1)},\, \mathbf{X}'_{(1)}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\,].$$

It can be found from Lemma 4(h) that

$$\begin{aligned} & r\left(\mathbf{X}' - \mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X}(\mathbf{X}'_{(1)}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{X}_{(1)})^{-1}[\,\mathbf{X}'_{(1)},\, \mathbf{X}'_{(1)}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\,]\right) \\ &= r[\,\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X} - \mathbf{X}'_{(1)}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{X}_{(1)},\, \mathbf{X}'_{(2)} - \mathbf{X}'_{(1)}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\,]. \end{aligned}$$

Hence (36) is equivalent to (c). The equivalence of (c) and (d) follows from Lemma 4(h). $\square$

The verification of the following theorem is left for the reader.

Theorem 15 *Let $\mathscr{M}_c$ and $\mathscr{M}_{1c}$ be as given in* (4) *and* (8), *and suppose* $r(\mathbf{X}_{(1)}) = p$ *and* $r(\boldsymbol{\Sigma}) = n$. *Then* $\mathrm{OLSE}_{\mathscr{M}_c}(\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}_{1c}}(\boldsymbol{\beta})$ *holds with probability 1 if and only if* $\mathscr{R}(\mathbf{X}'_{(2)}) \subseteq \mathscr{R}(\mathbf{A}')$.

Remarks. Many problems on relations between estimators for the general linear model can reasonably be investigated. For example, assume that the restricted linear model $\mathscr{M}_c$ in (3) incorrectly specifies the dispersion matrix $\boldsymbol{\Sigma}$ as $\sigma^2\boldsymbol{\Sigma}_0$, where Σ_0 is a given nonnegative definite matrix and σ^2 is a positive parameter (possibly unknown). Then consider relations between the BLUEs of the original model $\mathscr{M}_c$ and the misspecified model $\{\mathbf{y},\, \mathbf{X}\boldsymbol{\beta} \,|\, \mathbf{A}\boldsymbol{\beta} = \mathbf{b},\, \sigma^2\boldsymbol{\Sigma}_0\}$. Some results on relations between the BLUEs of two general linear models can be found in Mathew (1983), and Mitra and Moore (1973). Through the matrix rank method, we can also derive some necessary and sufficient conditions for the BLUEs of the original model and the misspecified model to equal with probability 1. This work will be presented in another paper.

In addition to the explicit restrictions in (2), the parameter vector $\boldsymbol{\beta}$ may be subject to some implicit restrictions, represented by adding-up conditions on the endogenous variables $\mathbf{B}\mathbf{y} = \mathbf{c}$, where both the $q \times n$ matrix $\mathbf{B}$ and $q \times 1$ vector $\mathbf{c}$ are fixed and known. Note that $\mathbf{y}$ is a random variable. Hence $\mathbf{B}\mathbf{y} = \mathbf{c}$ is equivalent to $E(\mathbf{B}\mathbf{y}) = \mathbf{c}$ and $Cov(\mathbf{B}\mathbf{y}) = \mathbf{0}$ with probability 1, i.e.,

$$\mathbf{B}\mathbf{X}\boldsymbol{\beta} = \mathbf{c} \ \text{ and } \mathbf{B}\boldsymbol{\Sigma} = \mathbf{0}. \tag{37}$$

In this case, (3), (4) and (37) can be written as

$$\mathscr{M}_r = \{\mathbf{y},\, \mathbf{X}\boldsymbol{\beta} \,|\, \mathbf{A}\boldsymbol{\beta} = \mathbf{b},\ \mathbf{B}\mathbf{X}\boldsymbol{\beta} = \mathbf{c},\ \mathbf{B}\boldsymbol{\Sigma} = \mathbf{0},\ \boldsymbol{\Sigma}\}.$$

This model is called the fully restricted linear model, see Haupt and Oberhofer (2002). It may also be of interest to investigate various equalities for the estimators of $\mathbf{X}\boldsymbol{\beta}$ in this model.

References

I.S. Alalouf and G.P.H. Styan (1979). Characterizations of estimability in the general linear model. *Ann. Statist.* 7, 194–200.

T. Amemiya (1985). Advanced Econometrics, Basil Blackwell, Oxford.

J.K. Baksalary and R.'Kala (1979). Best linear unbiased estimation in the restricted general linear model. *Math. Operationsforsch. Statist. Ser. Statist.* 10, 27–35.

J.K. Baksalary, S. Puntanen and G.P.H. Styan (1990). A property of the dispersion matrix of the best linear unbiased estimator in the general Gauss-Markov model. *Sankhyā Ser. A* 52, 279–296.

B.H. Baltagi (1989). Applications of a necessary and sufficient condition for OLS to be BLUE. *Statist. & Probab. Letters* 8, 457–461.

R. Bewley (1986). Allocation Models: Specifications, Estimation, and Application. Ballinger Publishing Company.

P. Bhimasankaram and R. SahaRay (1997). On a partitioned linear model and some associated reduced models. *Linear Algebra Appl.* 264, 329–339.

J.S. Chipman, M.M. Rao (1964). The treatment of linear restrictions in regression analysis. *Econometrica* 32, 198–209.

W.T. Dent (1980). On restricted estimation in linear models. J. Econometrics 12, 45–58.

C. Gourieroux and A. Monfort (1980). Sufficient linear structures: econometric applications. *Econometrica* 48, 1083–1097.

J. Groß and S. Puntanen (2000). Estimation under a general partitioned linear model. *Linear Algebra Appl.* 321, 131–144.

J. Groß, T. Trenkler and H.J. Werner (2001). The equality of linear transformations of the ordinary least squares estimator and the best linear unbiased estimator. *Sankhyā* 63, 118–127.

C.R. Hallum, T.O. Lewis and T.L. Boullion (1973). Estimation in the restricted general linear model with a positive semidefinite covariance matrix. *Comm. Statist.* 1, 157–166.

H. Haupt and W. Oberhofer (2002). Fully restricted linear regression: A pedagogical note. *Economics Bulletin* 3, 1–7.

G. Marsaglia and G.P.H. Styan (1974). Equalities and inequalities for ranks of matrices. *Linear and Multilinear Algebra* 2, 269–292.

T. Mathew (1983). A note on best linear unbiased estimation in the restricted general linear model. *Math. Operationsforsch. Statist. Ser. Statist.* 14, 3–6.

S.K. Mitra and B.J. Moore (1973). Gauss-Markov estimation with an incorrect dispersion matrix. *Sankhyā* Ser. A 35, 139–152.

M. Nurhonen and S. Puntanen (1992). A properties a partitioned generalized regression. *Commu. Statist. A-Theory Meth.* 21, 1579–1583.

S. Puntanen (1996). Some matrix results related to a partitioned singular linear model. *Commu. Statist.-Theory Meth.* 25, 269–279.

S. Puntanen and A.J. Scott (1996). Some further remarks on the singular linear models. *Linear Algebra Appl.* 237/238, 313–327.

S. Puntanen and G.P.H. Styan (1989). The equality of the ordinary least squares estimator and the best linear unbiased estimator. With comments by O. Kempthorne, S.R. Searle, and a reply by the authors. *Amer. Statist.* 43, 153–164.

S. Puntanen, G.P.H. Styan and H.J. Werner (2000). Two matrix-based proofs that the linear estimator Gy is the best linear unbiased estimator. *J. Statist. Plan. Infer.* 88, 173–179.

H. Qian and P. Schmidt (2003). Partial GLS regression. *Economic Letters* 79, 385–392.

H. Qian, Y. Tian (2006). Partially superfluous observations. *Econometric Theory* 22(2006), 529–536.

C.R. Rao (1971). Unified theory of linear estimation. *Sankhyā Ser. A* 33, 371–394.

C.R. Rao (1973). Representations of best linear unbiased estimators in the Gauss-Markoff model with a singular dispersion matrix. *J. Multivariate Anal.* 3, 276–292.

C.R. Rao and S.K. Mitra (1971). *Generalized Inverse of Matrices and Its Applications.* Wiley, New York.

B. Ravikumar, S. Ray and N.E. Savin (2000). Robust wald tests in SUR systems with adding up restrictions. *Econometrica* 68, 715–719.

S.R. Searle (1971). Linear Models. Wiley, New York.

H. Theil (1971). Principles of Econometrics. Wiley, New York.

Y. Tian and D.P. Wiens (2006). On equality and proportionality of ordinary least-squares, weighted least-squares and best linear unbiased estimators in the general linear model. *Statist. Prob. Lett.*, 76, 1265–1272.

H.J. Werner and C. Yapar (1995). More on partitioned possibly restricted linear regression. In *Proceedings of the 5th Tartu Conference on Multivariate Statistics* (E.-M. Tiit et al. eds.), 57–66.

W.-L. Yang, H.-J. Cui and G.-W. Sun (1987). On best linear unbiased estimation in the restricted general linear model. *Statistics* 18, 17–20.

B.-X. Zhang, B.-S. Liu and C.-Y. Lu (2004). A study of the equivalence of the BLUEs between a partitioned singular Linear model and its reduced singular linear models. *Acta Math. Sinica* 20, 557–568.