Projektbereich B
Discussion Paper No. B–301

More on partitioned possibly restricted linear regression

Hans Joachim WERNER*, Bonn, Germany
and
Cemil YAPAR, Trabzon, Turkey

1994

This paper appears in: *Proceedings of the 5th Tartu Conference on Multivariate Statistics (VSP Press, Utrecht, in press for publication in 1995, Ene-Margit Tiit et al., eds.).*

* Financial support by Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn, is gratefully acknowledged.

More on partitioned possibly restricted linear regression

H. J. Werner [1]
Institute for Econometrics and Operations Research
Econometrics Unit, University of Bonn
D-53113 Bonn, Germany

C. Yapar
Karadeniz Technical University
Faculty of Arts and Sciences, Dept. of Stat. & Comp. Sci.
61080 Trabzon, Turkey

Abstract

This paper deals with the *general* partitioned linear regression model where the regressor matrix $X = (\, X_1 \quad X_2 \,)$ may be deficient in column rank, the dispersion matrix V is possibly singular, $\beta^t = (\, \beta_1^t \quad \beta_2^t \,)$ - being partitioned according to X - is the vector of unknown regression coefficients, and β_2 is possibly subject to consistent linear equality or inequality restrictions. In particular, we are interested in the set of *generalized least squares (GLS) selections* for β_2. Inspired by Aigner and Balestra [1], as well as by Nurhonen and Puntanen [2], we also consider a specific reduced model and describe a scenario under which the set of GLS selections for β_2 under the reduced model equals the set of GLS selections for β_2 under the original full model. The results obtained in [2] and [1] for the unrestricted *standard* (full rank) regression model are reobtained as special cases.

Keywords: Gauß-Markov model, singular model, perfect multicollinearity, partitioned linear regression, linear equality constraints, linear inequality constraints, constrained generalized least squares selections, oblique projectors, generalized inverses.

JEL-Classification: C20.

---

[1] Financial support by Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn, is gratefully acknowledged.

More on partitioned possibly restricted linear regression

H. J. Werner [1]
Institute for Econometrics and Operations Research
Econometrics Unit, University of Bonn
D-53113 Bonn, Germany

C. Yapar
Karadeniz Technical University
Faculty of Arts and Sciences, Dept. of Stat. & Comp. Sci.
61080 Trabzon, Turkey

Abstract

This paper deals with the *general* partitioned linear regression model where the regressor matrix $X = (\, X_1 \quad X_2 \,)$ may be deficient in column rank, the dispersion matrix V is possibly singular, $\beta^t = (\, \beta_1^t \quad \beta_2^t \,)$ - being partitioned according to X - is the vector of unknown regression coefficients, and β_2 is possibly subject to consistent linear equality or inequality restrictions. In particular, we are interested in the set of *generalized least squares (GLS) selections* for β_2. Inspired by Aigner and Balestra [1], as well as by Nurhonen and Puntanen [2], we also consider a specific reduced model and describe a scenario under which the set of GLS selections for β_2 under the reduced model equals the set of GLS selections for β_2 under the original full model. The results obtained in [2] and [1] for the unrestricted *standard* (full rank) regression model are reobtained as special cases.

---

1. Introduction

Let $\mathbb{R}^n$, $\mathbb{R}^{n,m}$, and $\mathcal{P}^{n,n}$ denote the set of n-dimensional real column vectors, the set of $n \times m$ real matrices, and the set of real nonnegative definite (nnd) and symmetric $n \times n$ matrices, respectively. Given $A \in \mathbb{R}^{n,m}$, let A^t, $A^\dagger$, $\mathcal{R}(A)$, $\mathcal{N}(A)$, and $\mathrm{rank}(A)$ denote the transpose, the Moore-Penrose inverse, the range space (or column space), the null space, and the rank, respectively, of A. In addition, let A^- denote an arbitrary *g-inverse* of A satisfying $AA^-A = A$; denote the set of all g-inverses of A by $\{A^-\}$. Further, let I and 0, respectively, stand for the identity and zero matrix of whatever size is appropriate to the context.

[1] Financial support by Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn, is gratefully acknowledged.

Consider the Gauß-Markov model

$$\mathcal{L}(\mathcal{R}) := (y,\ X_1\beta_1 + X_2\beta_2,\ V \mid \beta \in \mathcal{R})$$

in which y is an observable random vector with expectation $X_1\beta_1 + X_2\beta_2$ and dispersion V; the vector $\beta := (\,\beta_1^t \quad \beta_2^t\,)^t$ of unknown regression coefficients satisfies the a priori restriction $\beta \in \mathcal{R}$; $\mathcal{R}$ is of one of the following cases:

- $\mathcal{R} = \mathcal{R}_e := \{\beta \mid R\beta = r\}$ (*e*quality constrained case),
- $\mathcal{R} = \mathcal{R}_i := \{\beta \mid R\beta \geq r\}$ (*i*nequality constrained case),
- $\mathcal{R} = \mathcal{R}_u := \mathbb{R}^m$, $m := m_1 + m_2$ (*u*nconstrained case);

and the model matrices X_1 $(n \times m_1)$, X_2 $(n \times m_2)$, V, R $(p \times m)$, $r \in \mathbb{R}^p$ are fixed and known. There are no rank assumptions on $X := (\,X_1 \quad X_2\,)$ and V. However, for $R = (\,R_1 \quad R_2\,)$ being partitioned according to $X = (\,X_1 \quad X_2\,)$ we assume throughout that $R_1 = 0$ and $\operatorname{rank}(R) = p$. In addition, we require that the conditions

$$\mathcal{R}(X) \subseteq \mathcal{R}(V) \tag{1.1}$$

and

$$\mathcal{R}(X_1) \cap \mathcal{R}(X_2) = \{0\} \tag{1.2}$$

are always satisfied.

For each $a \in \{e, i, u\}$: let $\mathcal{L}_a$ briefly denote the model $\mathcal{L}(\mathcal{R}_a)$. Note that, under the assumption (1.1),

$$y \in \mathcal{R}(V) \quad \text{(almost surely)}, \tag{1.3}$$

irrespective of under which of the three models y is observed. Due to invariance (cf. Theorem 2.3 in [4]), an arbitrary g-inverse V^- of V can be used to define a norm

$$\|x\|_{V^-} := (x^t V^- x)^{\frac{1}{2}}$$

on $\mathcal{R}(V)$. The mathematical programming problem

$$\text{minimize} \quad \|y - Xb\|_{V^-}^2 \quad \text{subject to} \quad b \in \mathcal{R}_a \tag{1.4}$$

is hence well defined for each realization $y \in \mathcal{R}(V)$. Any optimal solution to this convex-quadratic optimization program, that is, any vector from

$$\operatorname{argmin}_{b \in \mathcal{R}_a} \|y - Xb\|_{V^-}^2 \tag{1.5}$$

is called a *GLS solution* (for β) under model $\mathcal{L}_a$. Although (1.4) possesses an optimal solution for each $y \in \mathcal{R}(V)$, (1.5) need not be a singleton. In which case there do exist many different functions f with $f(y)$ representing a GLS solution for each $y \in \mathcal{R}(V)$. As in [4] we call any such function a *GLS selection* for β under $\mathcal{L}_a$; notice that it seems reasonable to reserve the term *'GLS estimator'* for exactly those situations where there does exist only one GLS selection on $\mathcal{R}(V)$. The set of all GLS selections for β under model $\mathcal{L}_a$ will be denoted by $\tilde{\beta}(\mathcal{L}_a)$.

Next, let $M_1 := I - X_1 X_1^\dagger$. Observe that M_1 coincides with the *orthogonal projector* (with respect to the standard inner product) onto $\mathcal{N}(X_1^t)$, that is, $M_1 = P_{\mathcal{N}(X_1^t)}$; cf. Section 2 in [4]. Hence, in particular, $M_1 X_1 = 0$, and for each $a \in \{e, i, u\}$ it is clear that model

$$\mathcal{M}_a := (M_1 y,\ M_1 X_2 \beta_2,\ M_1 V M_1 \mid \beta \in \mathcal{R}_a)$$

can be obtained from model $\mathcal{L}_a$ by premultiplying y by M_1. Besides these correctly *transformed* models $\mathcal{M}_a$, $a \in \{e, i, u\}$, the *reduced* models

$$\mathcal{N}_a := (M_1 y,\ M_1 X_2 \beta_2,\ V \mid \beta \in \mathcal{R}_a),\ a \in \{e, i, u\},$$

are also of interest to us in this paper. Unlike $\mathcal{M}_a$ in $\mathcal{N}_a$ the dispersion of $M_1 y$ is defined to be V. For each $a \in \{e, i, u\}$: let $\tilde{\beta}_2(\mathcal{M}_a)$ and $\tilde{\beta}_2(\mathcal{N}_a)$ stand for the set of all GLS selections for β_2 under $\mathcal{M}_a$ and under $\mathcal{N}_a$, respectively.

For convenience, we further introduce

$$J_1 := \begin{pmatrix} I \\ 0 \end{pmatrix} \quad \text{and} \quad J_2 := \begin{pmatrix} 0 \\ I \end{pmatrix}$$

such that $X_1 = X J_1$ and $X_2 = X J_2$. For each $a \in \{e, i, u\}$, we define $\tilde{\beta}_2(\mathcal{L}_a) := J_2^t \tilde{\beta}(\mathcal{L}_a)$; that is, $\tilde{\beta}_2(\mathcal{L}_a)$ denotes the set of all GLS selections for β_2 under the full model $\mathcal{L}_a$.

This paper is organized as follows. The main results are established in Section 3. There we prove that for each $a \in \{e, i, u\}$ we have $\tilde{\beta}_2(\mathcal{L}_a) = \tilde{\beta}_2(\mathcal{M}_a)$. In addition, we show that $\tilde{\beta}_2(\mathcal{L}_a) = \tilde{\beta}_2(\mathcal{N}_a)$ holds if $V\mathcal{R}(X_1) \subseteq \mathcal{R}(X_1)$ or, equivalently, $V\mathcal{N}(X_1^t) \subseteq \mathcal{N}(X_1^t)$ is satisfied. Recently (see [4]) we have derived some interesting representations relating the GLS selections of an equality or inequality constrained model to the GLS selections of the associated unconstrained model. In Section 2, some of these results, playing a key role in Section 3, are restated in the framework of our particular models.

2. Preliminary Results

For the sake of clarity we begin this section with quoting the following known result; see Theorem 2.4 and Theorem 2.10 (ii) in [4].

Theorem 2.1. For given $A \in \mathbb{R}^{n,m}$ and $W \in \mathcal{P}^{n,n}$, let $\mathcal{G}(A \mid W)$ denote the set of all those matrices $G \in \mathbb{R}^{m,n}$ which satisfy

$$G \in \{A^-\} \quad \text{and} \quad W\mathcal{N}(A^t) \subseteq \mathcal{N}(AG). \tag{2.1}$$

Then $\mathcal{G}(A \mid W)$ is nonempty. Moreover, if A is of full column rank, then GWG^t is nnd and invariant for any choice of $G \in \mathcal{G}(A \mid W)$.

Since $\mathcal{R}(X) \subseteq \mathcal{R}(V)$ [see (1.1)], Theorem 2.3 in [4] tells us that

$$\Gamma_{\mathcal{L}} := X^t V^- X \tag{2.2}$$

is nnd as well as invariant for any choice of V^-. Consider the restrictor matrix $R = (\,0 \quad R_2\,)$ from Section 1. Because $\Gamma_{\mathcal{L}} \in \mathcal{P}^{m,m}$, and since R is assumed to be of full row rank, it follows from

Theorem 2.1 that $(R^Q)^t\Gamma_{\mathcal{L}}R^Q$ is nnd and invariant for any choice of $(R^Q)^t \in \mathcal{G}(R^t \mid \Gamma_{\mathcal{L}})$; call this unique matrix $\Omega_{\mathcal{L}}$. Likewise it can be seen that

$$\Gamma_{\mathcal{M}} := (M_1X_2)^t(M_1VM_1)^-(M_1X_2) \in \mathcal{P}^{m_2,m_2} \tag{2.3}$$

and

$$\Gamma_{\mathcal{N}} := (M_1X_2)^tV^-(M_1X_2) \in \mathcal{P}^{m_2,m_2}, \tag{2.4}$$

are, respectively, invariant for any choice of $(M_1VM_1)^-$ and V^-. In other words, $\Gamma_{\mathcal{M}}$ and $\Gamma_{\mathcal{N}}$ are also well and uniquely defined nnd matrices. Since R_2 is of full row rank and as $\Gamma_{\mathcal{M}}$ is nnd, it again follows from Theorem 2.1 that $(R_2^Q)^t\Gamma_{\mathcal{M}}R_2^Q$ is nnd and invariant for any choice of $(R_2^Q)^t \in \mathcal{G}(R_2^t \mid \Gamma_{\mathcal{M}})$; call this matrix $\Omega_{\mathcal{M}}$. And finally, on similar lines we obtain that $(R_2^Q)^t\Gamma_{\mathcal{N}}R_2^Q$, henceforth denoted by $\Omega_{\mathcal{N}}$, is nnd and invariant for any choice of $(R_2^Q)^t \in \mathcal{G}(R_2^t \mid \Gamma_{\mathcal{N}})$.

Next, let $\mathcal{A} \in \{\mathcal{L}, \mathcal{M}, \mathcal{N}\}$. Choose any selection $\hat{\beta}_2 \in \tilde{\beta}_2(\mathcal{A}_u)$, and define $\hat{\mu}_{\mathcal{A}} := R_2\hat{\beta}_2 - r$. For each realization $y \in \mathcal{R}(V)$ (that is, for each realization $\hat{\mu}_{\mathcal{A}}$) it then follows from [4; see the lines around (3.18)] that

$$\tilde{M}_{\mathcal{A}}(y) := \operatorname{argmin}_{\mu \geq 0} \|\hat{\mu}_{\mathcal{A}} - \mu\|^2_{\Omega_{\mathcal{A}}} \tag{2.5}$$

is nonempty. Let $\tilde{M}_{\mathcal{A}}$ be pointwise defined according to (2.5). It is pertinent to mention here that $\tilde{M}_{\mathcal{A}}$ is invariant for any choice of $\hat{\beta}_2$. This happens although $\hat{\mu}_{\mathcal{A}}$, in general, does not share this property with $\tilde{M}_{\mathcal{A}}$; for more details, we refer to [4].

As announced in Section 1 we next state, for each $a \in \{e, i, u\}$, representations for $\tilde{\beta}(\mathcal{L}_a)$, $\tilde{\beta}_2(\mathcal{M}_a)$ and $\tilde{\beta}_2(\mathcal{N}_a)$. All these representations are readily obtained as special cases from Theorem 3.2, Theorem 3.3 and Theorem 3.4 in [4]; observe that (1.2) is equivalent to $\mathcal{N}(M_1X_2) = \mathcal{N}(X_2)$.

Theorem 2.2. Consider $\mathcal{L}_i$, $\mathcal{L}_e$, and $\mathcal{L}_u$. Let $\hat{\beta}_{\mathcal{L}}$ be an arbitrary but fixed GLS selection for β under model $\mathcal{L}_u$, and let $\Gamma_{\mathcal{L}}$ be defined as above. If $X^{\mathcal{L}}$ and $(R^Q)^t$ denote arbitrary but fixed matrices from $\mathcal{G}(X \mid V)$ and $\mathcal{G}(R^t \mid \Gamma_{\mathcal{L}})$, respectively, then we have

$$\tilde{\beta}(\mathcal{L}_i) = \{R^Q(r + \tilde{\mu}) + (I - R^QR)\hat{\beta}_{\mathcal{L}} + P_{\mathcal{N}(X)\cap\mathcal{N}(R)}w \mid \tilde{\mu} \in \tilde{M}_{\mathcal{L}},\ w \text{ arbitrary}\}, \tag{2.6}$$

$$\tilde{\beta}(\mathcal{L}_e) = \{R^Qr + (I - R^QR)\hat{\beta}_{\mathcal{L}} + P_{\mathcal{N}(X)\cap\mathcal{N}(R)}w \mid w \text{ arbitrary}\}, \tag{2.7}$$

$$\tilde{\beta}(\mathcal{L}_u) = \{X^{\mathcal{L}}y + P_{\mathcal{N}(X)}w \mid w \text{ arbitrary}\}, \tag{2.8}$$

where $\tilde{M}_{\mathcal{L}}$ is pointwise defined according to (2.5), and where $P_{\mathcal{N}(X)}$, for instance, denotes as usual the orthogonal projector onto $\mathcal{N}(X)$ [along $\mathcal{R}(X^t)$].

Theorem 2.3. Consider $\mathcal{M}_i$, $\mathcal{M}_e$, and $\mathcal{M}_u$. Let $\hat{\beta}_{2,\mathcal{M}}$ be an arbitrary but fixed GLS selection for β_2 under model $\mathcal{M}_u$, and let $\Gamma_{\mathcal{M}}$ be defined as above. If $(M_1X_2)^{\mathcal{M}}$ and $(R_2^Q)^t$ denote arbitrary but fixed matrices from $\mathcal{G}(M_1X_2 \mid M_1VM_1)$ and $\mathcal{G}(R_2^t \mid \Gamma_{\mathcal{M}})$, respectively, then we have

$$\tilde{\beta}_2(\mathcal{M}_i) = \{R_2^Q(r + \tilde{\mu}) + (I - R_2^QR_2)\hat{\beta}_{2,\mathcal{M}} + P_{\mathcal{N}(X_2)\cap\mathcal{N}(R_2)}w_2 \mid \tilde{\mu} \in \tilde{M}_{\mathcal{M}},\ w_2 \text{ arbitrary}\}, \tag{2.9}$$

$$\tilde{\beta}_2(\mathcal{M}_e) = \{R_2^Qr + (I - R_2^QR_2)\hat{\beta}_{2,\mathcal{M}} + P_{\mathcal{N}(X_2)\cap\mathcal{N}(R_2)}w_2 \mid w_2 \text{ arbitrary}\}, \tag{2.10}$$

$$\tilde{\beta}_2(\mathcal{M}_u) = \{(M_1X_2)^{\mathcal{M}}M_1y + P_{\mathcal{N}(X_2)}w_2 \mid w_2 \text{ arbitrary}\}, \tag{2.11}$$

where $\tilde{M}_{\mathcal{M}}$ is pointwise defined according to (2.5).

Theorem 2.4. Consider $\mathcal{N}_i$, $\mathcal{N}_e$, and $\mathcal{N}_u$. Let $\hat{\beta}_{2,\mathcal{N}}$ be an arbitrary but fixed GLS selection for β_2 under $\mathcal{N}_u$, and let $\Gamma_{\mathcal{N}}$ be defined as above. If $(M_1X_2)^{\mathcal{N}}$ and $(R_2^Q)^t$ denote arbitrary but fixed matrices from $\mathcal{G}(M_1X_2 \mid V)$ and $\mathcal{G}(R_2^t \mid \Gamma_{\mathcal{N}})$, respectively, then we have

$$\tilde{\beta}_2(\mathcal{N}_i) = \{R_2^Q(r+\tilde{\mu}) + (I - R_2^Q R_2)\hat{\beta}_{2,\mathcal{N}} + P_{\mathcal{N}(X_2)\cap\mathcal{N}(R_2)}w_2 \mid \tilde{\mu} \in \tilde{M}_{\mathcal{N}},\ w_2 \text{ arbitrary}\}, \quad (2.12)$$
$$\tilde{\beta}_2(\mathcal{N}_e) = \{R_2^Q r + (I - R_2^Q R_2)\hat{\beta}_{2,\mathcal{N}} + P_{\mathcal{N}(X_2)\cap\mathcal{N}(R_2)}w_2 \mid w_2 \text{ arbitrary}\}, \quad (2.13)$$
$$\tilde{\beta}_2(\mathcal{N}_u) = \{(M_1X_2)^{\mathcal{N}}M_1y + P_{\mathcal{N}(X_2)}w_2 \mid w_2 \text{ arbitrary}\}, \quad (2.14)$$

where $\tilde{M}_{\mathcal{N}}$ is pointwise defined according to (2.5).

3. Main Results

We formulate our main theorem as follows.

Theorem 3.1. For the linear models introduced in Section 1 we have:

(i) $\tilde{\beta}_2(\mathcal{L}_a) = \tilde{\beta}_2(\mathcal{M}_a)$ for each $a \in \{i,e,u\}$.

(ii) If $V\mathcal{R}(X_1) \subseteq \mathcal{R}(X_1)$, then $\tilde{\beta}_2(\mathcal{L}_a) = \tilde{\beta}_2(\mathcal{N}_a)$ for each $a \in \{i,e,u\}$.

For proving this theorem we need some further auxiliary results.

Theorem 3.2. Let (1.2) be satisfied for the block partitioned matrix $X = (\,X_1 \quad X_2\,) \in \mathbb{R}^{n,m}$, and let $V \in \mathcal{P}^{n,n}$. Moreover, let M_1, J_1, and J_2 be all defined as before. If $X^{\mathcal{L}} \in \mathcal{G}(X \mid V)$, then $J_2^tX^{\mathcal{L}} \in \mathcal{G}(M_1X_2 \mid M_1VM_1)$. In addition, $X_2J_2^tX^{\mathcal{L}}X_1 = 0$ or, equivalently, $J_2^tX^{\mathcal{L}}(I-M_1)z \in \mathcal{N}(X_2)$ for each $z \in \mathbb{R}^n$.

Proof. Let $X^{\mathcal{L}} \in \mathcal{G}(X \mid V)$. For $i = 1,2$, put $X_i^{\mathcal{L}} := J_i^tX^{\mathcal{L}}$. Then

$$X^{\mathcal{L}} = \begin{pmatrix} X_1^{\mathcal{L}} \\ X_2^{\mathcal{L}} \end{pmatrix}.$$

By definition of $\mathcal{G}(X \mid V)$, $XX^{\mathcal{L}}X = X$ and $V\mathcal{N}(X^t) \subseteq \mathcal{N}(XX^{\mathcal{L}})$. In view of (1.2), clearly $XX^{\mathcal{L}}X_1 = X_1$ and $XX^{\mathcal{L}}X_2 = X_2$ iff

$$X_iX_i^{\mathcal{L}}X_j = \begin{cases} X_i & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases} \quad (3.1)$$

for all $i,j \in \{1,2\}$. Therefore $X_2X_2^{\mathcal{L}}M_1 = X_2X_2^{\mathcal{L}}$ or, equivalently, $X_2X_2^{\mathcal{L}}(I-M_1) = 0$. This in turn implies $M_1X_2X_2^{\mathcal{L}}M_1X_2 = M_1X_2X_2^{\mathcal{L}}X_2 = M_1X_2$ so that we also have $X_2^{\mathcal{L}} \in \{(M_1X_2)^-\}$. It thus remains to show that $M_1VM_1\mathcal{N}(X_2^tM_1) \subseteq \mathcal{N}(M_1X_2X_2^{\mathcal{L}})$. Of course, $VM_1\mathcal{N}(X_2^tM_1) = V\mathcal{N}(X^t)$; for recall that M_1 represents the orthogonal projector onto $\mathcal{N}(X_1^t)$ along $\mathcal{R}(X_1)$. From (1.2) it further follows that

$$\mathcal{N}(XX^{\mathcal{L}}) = \mathcal{N}(X_1X_1^{\mathcal{L}}) \cap \mathcal{N}(X_2X_2^{\mathcal{L}}). \quad (3.2)$$

But then

$$M_1X_2X_2^{\mathcal{L}}(M_1VM_1)\mathcal{N}(X_2^tM_1) = M_1X_2X_2^{\mathcal{L}}V\mathcal{N}(X^t) = \{0\},$$

since $V\mathcal{N}(X^t) \subseteq \mathcal{N}(XX^{\mathcal{L}})$ and since $X_2X_2^{\mathcal{L}}M_1 = X_2X_2^{\mathcal{L}}$. This completes the proof. ∎

Theorem 3.3. Let $V \in \mathcal{P}^{n,n}$, and let (1.1) and (1.2) be satisfied for the block partitioned matrix $X = (\,X_1 \quad X_2\,) \in \mathbb{R}^{n,m}$. Further, let M_1, J_1, J_2, $\Gamma_{\mathcal{L}}$ and $\Gamma_{\mathcal{M}}$ be all defined as before. For convenience, put $\Gamma_{ij} := J_i^t \Gamma_{\mathcal{L}} J_j$, $i, j \in \{1, 2\}$. As usual, let $[\Gamma_{\mathcal{L}}/\Gamma_{11}]$ denote the *(generalized) Schur complement* of Γ_{11} in $\Gamma_{\mathcal{L}}$, that is, let $[\Gamma_{\mathcal{L}}/\Gamma_{11}] = \Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-}\Gamma_{12}$. Then $[\Gamma_{\mathcal{L}}/\Gamma_{11}] = \Gamma_{\mathcal{M}}$. Moreover, $J_2^t \Gamma_{\mathcal{L}}^{-} J_2 \in \{[\Gamma_{\mathcal{L}}/\Gamma_{11}]^{-}\}$, irrespective of $\Gamma_{\mathcal{L}}^{-}$. Hence, in particular, $J_2^t \Gamma_{\mathcal{L}}^{-} J_2 \in \{\Gamma_{\mathcal{M}}^{-}\}$.

Proof. First notice that (1.1) holds iff for $i = 1, 2$ we have $\mathcal{R}(X_i) \subseteq \mathcal{R}(V)$ or, equivalently, $\mathcal{N}(V) \subseteq \mathcal{N}(X_i^t)$. Next observe that $V^{-}V$ is a projector along $\mathcal{N}(V)$; cf. Section 2 in [4]. Consequently, $X_i^t V^{-} V = X_i^t$, for $i = 1, 2$. Since $\mathcal{R}(X_1) \subseteq \mathcal{R}(V)$, (2.6) in [4] gives us

$$\mathcal{R}(V) = \mathcal{R}(X_1) \oplus V\mathcal{N}(X_1^t)$$

with $\oplus$ indicating direct sum. In view of $\mathcal{R}(X_2) \subseteq \mathcal{R}(V)$, it should now be clear that there (uniquely) exist two matrices, say Z_1 and Z_2, such that $\mathcal{R}(Z_1) \subseteq \mathcal{R}(X_1)$, $\mathcal{R}(Z_2) \subseteq \mathcal{R}(VM_1) = V\mathcal{N}(X_1^t)$ and $X_2 = Z_1 + Z_2$. Since $P_1 := X_1(X_1^t V^{-} X_1)^{-} X_1^t V^{-}$ is a (generally oblique) projector onto $\mathcal{R}(X_1)$ with $V\mathcal{N}(X_1^t) \subseteq \mathcal{N}(P_1)$ (compare Theorem 2.5 in [4]), on the one hand $(I - P_1)X_2 = Z_2$. On the other hand we also get $VM_1(M_1VM_1)^{-}M_1X_2 = Z_2$ because $VM_1(M_1VM_1)^{-}M_1$ is a projector onto $\mathcal{R}(VM_1)$ satisfying $VM_1(M_1VM_1)^{-}M_1X_1 = 0$. Consequently $[I - X_1(X_1^tV^{-}X_1)^{-}X_1^tV^{-}]X_2 = VM_1(M_1VM_1)^{-}M_1X_2$. Combining all these observations results in

$$\begin{aligned}
[\Gamma_{\mathcal{L}}/\Gamma_{11}] &= \Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-}\Gamma_{12} \\
&= X_2^tV^{-}X_2 - X_2^tV^{-}X_1(X_1V^{-}X_1)^{-}X_1^tV^{-}X_2 \\
&= X_2^tV^{-}[I - X_1(X_1^tV^{-}X_1)^{-}X_1^tV^{-}]X_2 \\
&= X_2^tV^{-}VM_1(M_1VM_1)^{-}M_1X_2 \\
&= X_2^tM_1(M_1VM_1)^{-}M_1X_2, \\
&= \Gamma_{\mathcal{M}}.
\end{aligned}$$

As is well known (cf. [3], p. 46), $J_2^t\Gamma_{\mathcal{L}}^{-}J_2 \in \{[\Gamma_{\mathcal{L}}/\Gamma_{11}]^{-}\}$. Therefore $J_2^t\Gamma_{\mathcal{L}}^{-}J_2 \in \{\Gamma_{\mathcal{M}}^{-}\}$, and the proof is complete. ∎

Theorem 3.4. Consider the models $\mathcal{L}_i$ and $\mathcal{M}_i$, and let M_1, J_1, J_2, $\Gamma_{\mathcal{L}}$, $\Gamma_{\mathcal{M}}$, $\Omega_{\mathcal{L}}$ and $\Omega_{\mathcal{M}}$ be all defined as before. For each matrix $(R^Q)^t \in \mathcal{G}(R^t \mid \Gamma_{\mathcal{L}})$ we then have

$$(R^Q)^t J_2 \in \mathcal{G}(R_2^t \mid \Gamma_{\mathcal{M}}) \tag{3.3}$$

and

$$(R^Q)^t \Gamma_{\mathcal{L}} R^Q = (R^Q)^t J_2 \Gamma_{\mathcal{M}} J_2^t R^Q. \tag{3.4}$$

Therefore $\Omega_{\mathcal{L}} = \Omega_{\mathcal{M}}$.

Proof. Let $(R^Q)^t \in \mathcal{G}(R^t \mid \Gamma_{\mathcal{L}})$. Then, by definition of $\mathcal{G}(R^t \mid \Gamma_{\mathcal{L}})$,

$$R^Q \in \{R^{-}\} \quad \text{and} \quad (R^QR)^t\Gamma_{\mathcal{L}}\mathcal{N}(R) = \{0\}. \tag{3.5a-b}$$

Since R^t is of full column rank, (3.5 b) happens iff

$$(R^Q)^t\Gamma_{\mathcal{L}}\mathcal{N}(R) = \{0\}. \tag{3.6}$$

As $R = (\,0 \quad R_2\,)$, $\mathcal{N}(R) = \mathbb{R}^{m_1} \times \mathcal{N}(R_2)$ where '$\times$' indicates a cartesian product. But now it is evident that (3.6) is equivalent to

$$(R^Q)^t\Gamma_{\mathcal{L}}J_1 = 0, \quad (R^Q)^t\Gamma_{\mathcal{L}}J_2\mathcal{N}(R_2) = \{0\}. \tag{3.7a-b}$$

For convenience we put $\Gamma_{ij} := J_i^t \Gamma_{\mathcal{L}} J_j$ and $R_i^Q := J_i^t R^Q$ $(i, j = 1, 2)$. Then

$$\Gamma_{\mathcal{L}} = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \quad \text{and} \quad R^Q = \begin{pmatrix} R_1^Q \\ R_2^Q \end{pmatrix}.$$

By means of this notation, (3.7) can be rewritten as

$$(R_1^Q)^t \Gamma_{11} + (R_2^Q)^t \Gamma_{21} = 0, \quad [(R_1^Q)^t \Gamma_{12} + (R_2^Q)^t \Gamma_{22}] \mathcal{N}(R_2) = \{0\}. \tag{3.8a-b}$$

Since $\Gamma_{\mathcal{L}} := X^t V^- X$ is nnd, $\mathcal{R}(\Gamma_{12}) \subseteq \mathcal{R}(\Gamma_{11})$ [cf. [3], p. 71] or, equivalently, $\Gamma_{11}\Gamma_{11}^- \Gamma_{12} = \Gamma_{12}$. Postmultiplication of (3.8 a) by $\Gamma_{11}^- \Gamma_{12}$ therefore yields

$$(R_1^Q)^t \Gamma_{12} = -(R_2^Q)^t \Gamma_{21} \Gamma_{11}^- \Gamma_{12}. \tag{3.9}$$

Hence $(R_1^Q)^t \Gamma_{12} + (R_2^Q)^t \Gamma_{22} = (R_2^Q)^t [\Gamma_{\mathcal{L}}/\Gamma_{11}] = (R^Q)^t J_2 \Gamma_{\mathcal{M}}$ (observe Theorem 3.3), that is, we have

$$(R^Q)^t \Gamma_{\mathcal{L}} J_2 = (R^Q)^t J_2 \Gamma_{\mathcal{M}}. \tag{3.10}$$

(3.7 a) and (3.10) can be written as

$$(R^Q)^t \Gamma_{\mathcal{L}} = (\, 0 \quad (R^Q)^t J_2 \Gamma_{\mathcal{M}} \,), \tag{3.11}$$

whence we get

$$(R^Q)^t \Gamma_{\mathcal{L}} R^Q = (R^Q)^t J_2 \Gamma_{\mathcal{M}} J_2^t R^Q;$$

this is (3.4). In view of (3.11) and $\{0\} \times \mathcal{N}(R_2) \subseteq \mathcal{N}(R)$, we further get

$$(R^Q)^t J_2 \Gamma_{\mathcal{M}} \mathcal{N}(R_2) = \{0\}$$

directly from (3.6). Since $R = (\, 0 \quad R_2 \,)$, $R_2^Q \in \{R_2^-\}$ follows from (3.5 a). But now it is clear that (3.3) holds. From (3.3) and (3.4) we also get $\Omega_{\mathcal{L}} = \Omega_{\mathcal{M}}$; notice the lines directly following Theorem 2.1. This completes the proof. ∎

Theorem 3.5. Consider the linear models introduced in Section 1, and let

$$V\mathcal{R}(X_1) \subseteq \mathcal{R}(X_1) \quad \text{or, equivalently,} \quad V\mathcal{N}(X_1^t) \subseteq \mathcal{N}(X_1^t)$$

be satisfied. Moreover, let M_1, J_1, J_2, $\Gamma_{\mathcal{M}}$, $\Gamma_{\mathcal{N}}$, $\Omega_{\mathcal{M}}$ and $\Omega_{\mathcal{N}}$ be all defined as before. For each $X^{\mathcal{L}} \in \mathcal{G}(X \mid V)$, we then have $J_2^t X^{\mathcal{L}} \in \mathcal{G}(M_1 X_2 \mid V)$. In addition, we get $\Gamma_{\mathcal{N}} = \Gamma_{\mathcal{M}}$, and therefore $\mathcal{G}(R_2^t \mid \Gamma_{\mathcal{N}}) = \mathcal{G}(R_2^t \mid \Gamma_{\mathcal{M}})$ as well as $\Omega_{\mathcal{N}} = \Omega_{\mathcal{M}}$.

Proof. Because $M_1 = P_{\mathcal{N}(X_1^t)}$ and since $V\mathcal{N}(X_1^t) \subseteq \mathcal{N}(X_1^t)$, clearly $M_1 V M_1 = V M_1$. Since $(M_1 V M_1)^t = M_1 V M_1$, therefore $M_1 V M_1 = V M_1 = M_1 V$. Consequently

$$(M_1 V M_1) V^- (M_1 V M_1) = M_1 V V^- V M_1 = M_1 V M_1,$$

that is, $\{V^-\} \subseteq \{(M_1 V M_1)^-\}$. But now trivially $\Gamma_{\mathcal{M}} = \Gamma_{\mathcal{N}}$ and $\mathcal{G}(R_2^t \mid \Gamma_{\mathcal{M}}) = \mathcal{G}(R_2^t \mid \Gamma_{\mathcal{N}})$. Next, let $X^{\mathcal{L}} \in \mathcal{G}(X \mid V)$ be arbitrary but fixed. Put $X_i^{\mathcal{L}} := J_i^t X^{\mathcal{L}}$ for $i = 1, 2$. Then

$$X^{\mathcal{L}} = \begin{pmatrix} X_1^{\mathcal{L}} \\ X_2^{\mathcal{L}} \end{pmatrix}.$$

By Theorem 3.2, $X_2^{\mathcal{L}} \in \mathcal{G}(M_1X_2 \mid M_1VM_1)$. Hence, in particular, $X_2^{\mathcal{L}} \in \{(M_1X_2)^-\}$. So it remains to show that $M_1X_2X_2^{\mathcal{L}}V\mathcal{N}(X_2^tM_1) = \{0\}$ is also satisfied. Clearly, $V\mathcal{N}(X_2^tM_1) = V[\mathcal{R}(X_1) \oplus \mathcal{N}(X^t)] = V\mathcal{R}(X_1) + V\mathcal{N}(X^t)$. Since $V\mathcal{R}(X_1) \subseteq \mathcal{R}(X_1)$, it follows in view of (3.1) that $V\mathcal{R}(X_1) \subseteq \mathcal{N}(M_1X_2X_2^{\mathcal{L}})$. In the proof of Theorem 3.2 we have already seen that $V\mathcal{N}(X^t) \subseteq \mathcal{N}(M_1X_2X_2^{\mathcal{L}})$. This completes the proof. ∎

Now we are in the position to prove Theorem 3.1

Proof of Theorem 3.1 (i). First, we wish to establish $\tilde{\beta}_2(\mathcal{L}_u) = \tilde{\beta}_2(\mathcal{M}_u)$, that is, $J_2^t\tilde{\beta}(\mathcal{L}_u) = \tilde{\beta}_2(\mathcal{M}_u)$ where J_2 is defined as before. To that end, consider representation (2.8) for $\tilde{\beta}(\mathcal{L}_u)$ with an arbitrary but fixed $X^{\mathcal{L}} \in \mathcal{G}(X \mid V)$. Observe that it suffices to show that $J_2^t\tilde{\beta}(\mathcal{L}_u)$ is of the form (2.11). As (1.2) is equivalent to $\mathcal{N}(X) = \mathcal{N}(X_1) \times \mathcal{N}(X_2)$, clearly

$$P_{\mathcal{N}(X)} = \begin{pmatrix} P_{\mathcal{N}(X_1)} & 0 \\ 0 & P_{\mathcal{N}(X_2)} \end{pmatrix}.$$

If we now partition $w^t = (\, w_1^t \quad w_2^t \,)$ appropriately, then $J_2^tP_{\mathcal{N}(X)}w = P_{\mathcal{N}(X_2)}w_2$. From Theorem 3.2 we further know that $J_2^tX^{\mathcal{L}} \in \mathcal{G}(M_1X_2 \mid M_1VM_1)$ and that (for each possible realization y) $J_2^tX^{\mathcal{L}}y$ can be written alternatively in the form $J_2^tX^{\mathcal{L}}y = J_2^tX^{\mathcal{L}}M_1y + z$ for some suitably chosen $z \in \mathcal{N}(X_2)$. With these observations in mind, the claimed result is obvious.

Next, we prove $\tilde{\beta}_2(\mathcal{L}_e) = \tilde{\beta}_2(\mathcal{M}_e)$. Consider representation (2.7) for $\tilde{\beta}(\mathcal{L}_e)$ with an arbitrary but fixed $(R^Q)^t \in \mathcal{G}(R^t \mid \Gamma_{\mathcal{L}})$. In view of $R = (\, 0 \quad R_2 \,)$ and (1.2), clearly $\mathcal{N}(X) \cap \mathcal{N}(R) = \mathcal{N}(X_1) \times [\mathcal{N}(X_2) \cap \mathcal{N}(R_2)]$. As above we therefore get $J_2^tP_{\mathcal{N}(X)\cap\mathcal{N}(R)}w = P_{\mathcal{N}(X_2)\cap\mathcal{N}(R_2)}w_2$ where $w_2 = J_2^tw$. Theorem 3.4 tells us that $(J_2^t(R^Q))^t \in \mathcal{G}(R_2^t \mid \Gamma_{\mathcal{M}})$. In addition, we get $J_2^t(I - R^QR)\hat{\beta}_{\mathcal{L}} = (I - J_2^tR^QR_2)J_2^t\hat{\beta}_{\mathcal{L}}$ as $R = (\, 0 \quad R_2 \,)$. From the previous step we know that $J_2^t\hat{\beta}_{\mathcal{L}} \in \tilde{\beta}_2(\mathcal{M}_u)$. In view of all these observations it is now clear that $J_2^t\tilde{\beta}(\mathcal{L}_e)$ is a representation for $\tilde{\beta}(\mathcal{M}_e)$; see (2.10).

In order to prove $\tilde{\beta}_2(\mathcal{L}_i) = \tilde{\beta}_2(\mathcal{M}_i)$, consider representation (2.6) for $\tilde{\beta}(\mathcal{L}_i)$. In the light of (2.6), (2.7), (2.9), (2.10) and the foregoing step it suffices to show that $\tilde{M}_{\mathcal{L}} = \tilde{M}_{\mathcal{M}}$. But by Theorem 3.4, $\Omega_{\mathcal{L}} = \Omega_{\mathcal{M}}$. Hence $\tilde{M}_{\mathcal{L}} = \tilde{M}_{\mathcal{M}}$ as $\tilde{\beta}_2(\mathcal{L}_u) = \tilde{\beta}_2(\mathcal{M}_u)$. ∎

Proof of Theorem 3.1 (ii). Let $V\mathcal{R}(X_1) \subseteq \mathcal{R}(X_1)$. Then, by Theorem 3.5, $J_2^tX^{\mathcal{L}} \in \mathcal{G}(M_1X_2 \mid V)$ for each $X^{\mathcal{L}} \in \mathcal{G}(X \mid V)$. By the same theorem, $\mathcal{G}(R_2^t \mid \Gamma_{\mathcal{N}}) = \mathcal{G}(R_2^t \mid \Gamma_{\mathcal{M}})$ Therefore, in view of (3.3), $(R^Q)^tJ_2 \in \mathcal{G}(R_2^t \mid \Gamma_{\mathcal{N}})$ for each $(R^Q)^t \in \mathcal{G}(R^t \mid \Gamma_{\mathcal{L}})$. As in the proof of part (i), we therefore get $\tilde{\beta}_2(\mathcal{L}_u) := J_2^t\tilde{\beta}(\mathcal{L}_u) = \tilde{\beta}_2(\mathcal{N}_u)$ and $\tilde{\beta}_2(\mathcal{L}_e) := J_2^t\tilde{\beta}(\mathcal{L}_e) = \tilde{\beta}_2(\mathcal{N}_e)$ by means of Theorem 2.2 and Theorem 2.4. By Theorem 3.5, $\Omega_{\mathcal{M}} = \Omega_{\mathcal{N}}$. Since $\tilde{\beta}_2(\mathcal{L}_u) = \tilde{\beta}_2(\mathcal{M}_u) = \tilde{\beta}_2(\mathcal{N}_u)$, therefore $\tilde{M}_{\mathcal{M}} = \tilde{M}_{\mathcal{N}}$. But then $\tilde{\beta}_2(\mathcal{L}_i) := J_2^t\tilde{\beta}(\mathcal{L}_i) = \tilde{\beta}_2(\mathcal{N}_i)$, and the proof is complete. ∎

We conclude this paper with mentioning that if X and V are both of full column rank then all the GLS selections considered in this paper are unique, that is, the corresponding GLS estimators do exist. Notice that in such a situation the GLS estimator for β_2 under each of our models is the BLUE for β_2 in that model. The unrestricted full rank case (that is, the case where we have no a priori restrictions and where X and V are both of full column rank) has already been investigated recently in [2] and [1, p. 970].

References

[1] D. J. Aigner and P. Balestra. *Optimal experimental design for error components models.* Econometrica, 56: 955–971 (1988).

[2] M. Nurhonen and S. Puntanen. *A property of partitioned generalized regression.* Commun. Statist.-Theory Meth., 21(6): 1579–1583 (1992).

[3] R. M. Pringle and A. A. Rayner. *Generalized Inverse Matrices with Applications to Statistics.* Griffin, London, 1971.

[4] H. J. Werner and C. Yapar. *On inequality constrained generalized least squares selections in the general possibly singular Gauß-Markov model: a projector theoretical approach.* Linear Algebra Appl. (1995, to appear; Fifth Special Issue devoted to Linear Algebra and Statistics).