Econometric Theory, **22**, 2006, 529–536. Printed in the United States of America.
DOI: 10.1017/S0266466606060269

NOTES AND PROBLEMS

PARTIALLY SUPERFLUOUS OBSERVATIONS

Hailong Qian
Saint Louis University

Yongge Tian
Shanghai University of Finance and Economics

Necessary and sufficient conditions are established for the second subsample to be partially redundant given the first subsample for the best linear unbiased estimator of a subset of the coefficient vector of a general linear regression model.

1. PROBLEM AND MOTIVATION

Consider the general linear (Gauss–Markov) regression model

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i = \mathbf{x}_{i1}'\boldsymbol{\beta}_1 + \mathbf{x}_{i2}'\boldsymbol{\beta}_2 + \varepsilon_i, \qquad i = 1,\dots,N, \tag{1}$$

where N is the sample size, y_i is the dependent variable, $\mathbf{x}_i \equiv (\mathbf{x}_{i1}', \mathbf{x}_{i2}')'$ is a $K \times 1$ column vector of explanatory variables (with $\mathbf{x}_{i1}$ being $K_1 \times 1$, $\mathbf{x}_{i2}$ being $K_2 \times 1$, and $K_1 + K_2 = K$), $\boldsymbol{\beta} \equiv (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')'$ is a $K \times 1$ vector of coefficients that is partitioned accordingly, and $\boldsymbol{\varepsilon}_i$ is the disturbance term. Suppose that the sample consists of two subsamples; the first one is of size n and the second one is of size m, with $n + m = N$. The two subsamples can be from different sources, for example, from different surveys. We make the following assumptions.

(A1) $E(\boldsymbol{\varepsilon}|\mathbf{x}_1,\dots,\mathbf{x}_N) = \mathbf{0}$ with $\boldsymbol{\varepsilon} = (\varepsilon_1,\dots,\varepsilon_N)'$.

(A2) $\mathbf{X}_{1\bullet} = (\mathbf{x}_1,\dots,\mathbf{x}_n)'$ is of full column rank.

(A3) $\text{var}(\boldsymbol{\varepsilon}|\mathbf{x}_1,\dots,\mathbf{x}_N) = \sigma^2\mathbf{V}$, where σ^2 is a positive scalar and $\mathbf{V}$ is an $N \times N$ known positive definite matrix.

Under these assumptions, it is well known that the generalized least squares (GLS) estimator $\tilde{\boldsymbol{\beta}}$ using the entire sample is the best linear unbiased estimator

We thank the Co-editor for his very helpful comments on an earlier draft. Address correspondence to Hailong Qian, Department of Economics, Saint Louis University, 3674 Lindell Blvd., St. Louis, MO 63108, USA; e-mail: qianh@slu.edu.

 0266-4666/06 $12.00

(BLUE) of $\boldsymbol{\beta}$ and is thus more efficient than the GLS estimator $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ using the first subsample only; see, e.g., Gourieroux and Monfort (1980, p. 1093). When $\tilde{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}$ are equally efficient (in other words, when $\hat{\boldsymbol{\beta}}$ is also BLUE), Gourieroux and Monfort (1980) refer to the second subsample as superfluous observations for the efficient estimation of $\boldsymbol{\beta}$. In this note, we wish to extend their result to a subset of the regression coefficients; more specifically, we seek to find the necessary and sufficient conditions for the second subsample to be partially superfluous for the BLUE of $\boldsymbol{\beta}_1$ alone.

2. SOLUTION AND DISCUSSION

Stacking model (1) over the observations in the first subsample, we have

$$\mathbf{y}_1 = \mathbf{X}_{1\bullet}\boldsymbol{\beta} + \boldsymbol{\varepsilon}_1 = \mathbf{X}_{11}\boldsymbol{\beta}_1 + \mathbf{X}_{12}\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}_1, \tag{2a}$$

where

$$\mathbf{y}_1 = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \qquad \mathbf{X}_{11} = \begin{pmatrix} \mathbf{x}'_{11} \\ \vdots \\ \mathbf{x}'_{n1} \end{pmatrix}, \qquad \mathbf{X}_{12} = \begin{pmatrix} \mathbf{x}'_{12} \\ \vdots \\ \mathbf{x}'_{n2} \end{pmatrix},$$

$$\mathbf{X}_{1\bullet} = (\mathbf{X}_{11}, \mathbf{X}_{12}), \qquad \boldsymbol{\varepsilon}_1 = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Similarly, stacking model (1) over the observations in the second subsample, we have

$$\mathbf{y}_2 = \mathbf{X}_{2\bullet}\boldsymbol{\beta} + \boldsymbol{\varepsilon}_2 = \mathbf{X}_{21}\boldsymbol{\beta}_1 + \mathbf{X}_{22}\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}_2, \tag{2b}$$

where the notation is defined similarly to before. Finally, stacking (2a) over (2b), we have

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}, \tag{3}$$

where

$$\mathbf{y}_1 = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}, \qquad \mathbf{X}_1 = \begin{pmatrix} \mathbf{X}_{11} \\ \mathbf{X}_{21} \end{pmatrix}, \qquad \mathbf{X}_2 = \begin{pmatrix} \mathbf{X}_{12} \\ \mathbf{X}_{22} \end{pmatrix},$$

$$\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2) = \begin{pmatrix} \mathbf{X}_{1\bullet} \\ \mathbf{X}_{2\bullet} \end{pmatrix}, \qquad \boldsymbol{\varepsilon} = \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{pmatrix}.$$

Under assumptions (A1)–(A3), we can estimate $\boldsymbol{\beta}$ in (3) (or (1)) by GLS using the first subsample alone or using the entire sample. Thus, we define the following two GLS estimators of $\boldsymbol{\beta}$:

https://doi.org/10.1017/S0266466606060269 Published online by Cambridge University Press

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}_{1\bullet}'\mathbf{V}_{11}^{-1}\mathbf{X}_{1\bullet})^{-1}\mathbf{X}_{1\bullet}'\mathbf{V}_{11}^{-1}\mathbf{y}_1, \tag{4a}$$

$$\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}, \tag{4b}$$

where $\mathbf{V}_{11} = \text{var}(\boldsymbol{\varepsilon}_1)$.

Gourieroux and Monfort (1980, p. 1093) derived the necessary and sufficient condition for $\hat{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\beta}}$ to have the same finite-sample efficiency (i.e., the second subsample is fully superfluous for the efficient estimation of $\boldsymbol{\beta}$):

$$\mathbf{X}_{2\bullet} = \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{1\bullet}. \tag{5}$$

In this note, we wish to extend their result to a subvector of $\boldsymbol{\beta}$. More specifically, we seek to find necessary and sufficient conditions for the second subsample to be partially superfluous for the BLUE of $\boldsymbol{\beta}_1$ alone.

Under assumptions (A1)–(A3), it is well known that $\tilde{\boldsymbol{\beta}}$ is BLUE with $\text{var}(\tilde{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$ and $\hat{\boldsymbol{\beta}}$ is also unbiased with $\text{var}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}_{1\bullet}'\mathbf{V}_{11}^{-1}\mathbf{X}_{1\bullet})^{-1}$. Notice that $\boldsymbol{\beta}_1 = \mathbf{S}\boldsymbol{\beta}$, where the selection matrix $\mathbf{S}$ is defined as $\mathbf{S} = (\mathbf{I}_{K_1}, \mathbf{0}_{K_1\times K_2})$. We can then easily obtain

$$\text{var}(\hat{\boldsymbol{\beta}}_1) = \sigma^2\mathbf{S}(\mathbf{X}_{1\bullet}'\mathbf{V}_{11}^{-1}\mathbf{X}_{1\bullet})^{-1}\mathbf{S}', \tag{6a}$$

$$\text{var}(\tilde{\boldsymbol{\beta}}_1) = \sigma^2\mathbf{S}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{S}'. \tag{6b}$$

Because $\tilde{\boldsymbol{\beta}}$ is the BLUE for $\boldsymbol{\beta}$ in (1), $\text{var}(\hat{\boldsymbol{\beta}}_1) - \text{var}(\tilde{\boldsymbol{\beta}}_1)$ is always positive semidefinite and equals zero, if, and only if, $\text{var}(\hat{\boldsymbol{\beta}}_1) = \text{var}(\tilde{\boldsymbol{\beta}}_1)$, or equivalently,

$$\mathbf{S}(\mathbf{X}_{1\bullet}'\mathbf{V}_{11}^{-1}\mathbf{X}_{1\bullet})^{-1}\mathbf{S}' - \mathbf{S}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{S}' = \mathbf{0}. \tag{7}$$

There are several ways to find the necessary and sufficient conditions for equation (7) to be true; for example, one obvious but "brute-force" approach is to use the partitioned matrix inverse rule for $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$ and then simplify the resulting expression. In this note, we will use a much simpler and relatively novel approach to finding the necessary and sufficient conditions for (7) to hold; more specifically, we will calculate the rank of the matrix on the left-hand side of (7). For this purpose, we first recall a simple matrix fact: a matrix is a zero matrix if, and only if, its rank is zero. Thus, (7) holds if, and only if,

$$\text{rk}[\mathbf{S}(\mathbf{X}_{1\bullet}'\mathbf{V}_{11}^{-1}\mathbf{X}_{1\bullet})^{-1}\mathbf{S}' - \mathbf{S}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{S}'] = 0. \tag{8}$$

To find the rank of the left-hand side of (8), we need to use two well-known rank formulas for partitioned matrices due to Marsaglia and Styan (1974).

LEMMA 1. *(i) Let* $\mathbf{A}$ *be a nonsingular matrix of order* m*, let* $\mathbf{B}$*,* $\mathbf{C}$*, and* $\mathbf{D}$ *be* $m \times k$*,* $l \times m$*, and* $l \times k$ *matrices, respectively. Then*

$$\text{rk}\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \text{rk}\begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{pmatrix} = \text{rk}(\mathbf{A}) + \text{rk}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}). \tag{9}$$

https://doi.org/10.1017/S0266466606060269 Published online by Cambridge University Press

(ii) If $\mathbf{M}$ *is symmetric positive semidefinite, then*

$$\mathrm{rk}\begin{pmatrix} \mathbf{M} & \mathbf{B} \\ \mathbf{B}' & \mathbf{0} \end{pmatrix} = \mathrm{rk}(\mathbf{M}, \mathbf{B}) + \mathrm{rk}(\mathbf{B}). \tag{10}$$

It can now be easily seen from (9) that $\mathbf{CA}^{-1}\mathbf{B} = \mathbf{D}$ if and only if $\mathrm{rk}(\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}) = 0$, or equivalently $\mathrm{rk}\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \mathrm{rk}(\mathbf{A})$. In fact, any matrix expression involving inverse of matrix can always be rewritten in the form of the Schur complement $\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}$. Then, using (9), we can construct the corresponding partitioned matrix consisting of submatrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$. When the resulting partitioned matrix can be simplified by elementary block matrix operations, we can obtain some nontrivial rank formula for $\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}$, and, as a consequence, we can further derive necessary and sufficient conditions for $\mathbf{CA}^{-1}\mathbf{B} = \mathbf{D}$ to be true.

Given Lemma 1, we are now ready to derive the necessary and sufficient conditions for the second subsample to be partially superfluous for the BLUE of $\boldsymbol{\beta}_1$. We first establish Lemma 2, as follows.

LEMMA 2. *Partition the covariance matrix* $\sigma^2\mathbf{V}$ *into a* 2×2 *block matrix, with the* (i,j)*-block* $\mathbf{V}_{ij} = E(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_j')$ *for* $i,j = 1,2$. *Then,*

$$\begin{aligned} &\mathrm{rk}[\mathrm{var}(\hat{\boldsymbol{\beta}}_1) - \mathrm{var}(\tilde{\boldsymbol{\beta}}_1)] \\ &\quad = \mathrm{rk}[(\mathbf{X}_{21} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{11}) - (\mathbf{X}_{22} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{12}) \\ &\qquad \times (\mathbf{X}_{12}'\mathbf{V}_{11}^{-1}\mathbf{X}_{12})^{-1}\mathbf{X}_{12}'\mathbf{V}_{11}^{-1}\mathbf{X}_{11}]. \end{aligned} \tag{11}$$

Proof. See the Appendix.

Now, using Lemma 2 and the fact that a matrix is a zero matrix if and only if its rank is also equal to zero, we can readily obtain the main result of this note.

THEOREM 3. *Under assumptions (A1)–(A3), given the first subsample, the second subsample is partially superfluous for the BLUE of* $\boldsymbol{\beta}_1$ *in model (1) if, and only if,*

$$\mathbf{X}_{21} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{11} = (\mathbf{X}_{22} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{12})(\mathbf{X}_{12}'\mathbf{V}_{11}^{-1}\mathbf{X}_{12})^{-1}\mathbf{X}_{12}'\mathbf{V}_{11}^{-1}\mathbf{X}_{11}. \tag{12}$$

One way to provide an intuition for the necessary and sufficient condition (12) is as follows. Premultiplying (3) by $\mathbf{A}_1 = \begin{pmatrix} \mathbf{V}_{11}^{-1/2} & \mathbf{0} \\ -\mathbf{V}_{21}\mathbf{V}_{11}^{-1} & \mathbf{I}_m \end{pmatrix}$, we have

$$\begin{aligned} \begin{pmatrix} \mathbf{V}_{11}^{-1/2}\mathbf{y}_1 \\ \mathbf{y}_2 - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{y}_1 \end{pmatrix} &= \begin{pmatrix} \mathbf{V}_{11}^{-1/2}\mathbf{X}_{11} \\ \mathbf{X}_{21} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{11} \end{pmatrix}\boldsymbol{\beta}_1 \\ &\quad + \begin{pmatrix} \mathbf{V}_{11}^{-1/2}\mathbf{X}_{12} \\ \mathbf{X}_{22} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{12} \end{pmatrix}\boldsymbol{\beta}_2 + \begin{pmatrix} \mathbf{V}_{11}^{-1/2}\boldsymbol{\varepsilon}_1 \\ \mathbf{e}_2 \end{pmatrix} \end{aligned} \tag{13a}$$

https://doi.org/10.1017/S0266466606060269 Published online by Cambridge University Press

or simply written as

$$\begin{pmatrix} \mathbf{y}_1^* \\ \mathbf{y}_2^* \end{pmatrix} = \begin{pmatrix} \mathbf{X}_{11}^* \\ \mathbf{X}_{21}^* \end{pmatrix} \boldsymbol{\beta}_1 + \begin{pmatrix} \mathbf{X}_{12}^* \\ \mathbf{X}_{22}^* \end{pmatrix} \boldsymbol{\beta}_2 + \begin{pmatrix} \boldsymbol{\varepsilon}_1^* \\ \mathbf{e}_2 \end{pmatrix}, \tag{13b}$$

where $\mathbf{e}_2 = \boldsymbol{\varepsilon}_2 - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\boldsymbol{\varepsilon}_1$ is the linear projection error of $\boldsymbol{\varepsilon}_2$ onto $\boldsymbol{\varepsilon}_1$ and thus is uncorrelated with $\boldsymbol{\varepsilon}_1$. Now, premultiplying (13b) again by $\mathbf{A}_2 = \begin{pmatrix} \mathbf{M}_2 & \mathbf{0} \\ -\mathbf{X}_{22}^*(\mathbf{X}_{12}^{*\prime}\mathbf{X}_{12}^*)^{-1}\mathbf{X}_{12}^{*\prime} & \mathbf{I}_m \end{pmatrix}$ with $\mathbf{M}_2 = \mathbf{I}_n - \mathbf{X}_{12}^*(\mathbf{X}_{12}^{*\prime}\mathbf{X}_{12}^*)^{-1}\mathbf{X}_{12}^{*\prime}$, we have

$$\begin{pmatrix} \mathbf{M}_2\mathbf{y}_1^* \\ \mathbf{y}_2^{**} \end{pmatrix} = \begin{pmatrix} \mathbf{M}_2\mathbf{X}_{11}^* \\ \mathbf{X}_{21}^{**} \end{pmatrix} \boldsymbol{\beta}_1 + \begin{pmatrix} \mathbf{M}_2\boldsymbol{\varepsilon}_1^* \\ \mathbf{e}_2^{**} \end{pmatrix}, \tag{14}$$

where $\mathbf{y}_2^{**} = \mathbf{y}_2^* - \mathbf{X}_{22}^*(\mathbf{X}_{12}^{*\prime}\mathbf{X}_{12}^*)^{-1}\mathbf{X}_{12}^{*\prime}\mathbf{y}_1^*$, $\mathbf{X}_{21}^{**} = \mathbf{X}_{21}^* - \mathbf{X}_{22}^*(\mathbf{X}_{12}^{*\prime}\mathbf{X}_{12}^*)^{-1}\mathbf{X}_{12}^{*\prime}\mathbf{X}_{11}^*$, and $\mathbf{e}_2^{**} = \mathbf{e}_2 - \mathbf{X}_{22}^*(\mathbf{X}_{12}^{*\prime}\mathbf{X}_{12}^*)^{-1}\mathbf{X}_{12}^{*\prime}\boldsymbol{\varepsilon}_1^*$. Here we particularly notice that the two regression errors in (14), $\mathbf{M}_2\boldsymbol{\varepsilon}_1^*$ and $\mathbf{e}_2^{**}$, are uncorrelated as a result of $\text{cov}(\mathbf{e}_2, \boldsymbol{\varepsilon}_1^*) = \mathbf{0}$, $\text{var}(\boldsymbol{\varepsilon}_1^*) = \sigma^2\mathbf{I}_n$, and $\mathbf{X}_{12}^{*\prime}\mathbf{M}_2 = \mathbf{0}$. Then, if the error vector of the original regression model (3), $(\boldsymbol{\varepsilon}_1', \boldsymbol{\varepsilon}_2')'$, is distributed as multivariate normal, the conditional distribution of $\mathbf{y}_2^{**}$ given $\mathbf{y}_1$ and $\mathbf{X}$ is also multivariate normal, its covariance matrix does not depend on $\boldsymbol{\beta}_1$, and its conditional mean is $\mathbf{X}_{21}^{**}\boldsymbol{\beta}_1$. Thus, by the definition of $\mathbf{y}_2^{**} = (\mathbf{y}_2 - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{y}_1) - \mathbf{X}_{22}^*(\mathbf{X}_{12}^{*\prime}\mathbf{X}_{12}^*)^{-1}\mathbf{X}_{12}^{*\prime}\mathbf{V}_{11}^{-1/2}\mathbf{y}_1$, we can deduce that the conditional distribution of $\mathbf{y}_2$ given $\mathbf{y}_1$ and $\mathbf{X}$ is also normal, its covariance matrix does not depend on $\boldsymbol{\beta}_1$, and its conditional mean is $\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{y}_1 + \mathbf{X}_{22}^*(\mathbf{X}_{12}^{*\prime}\mathbf{X}_{12}^*)^{-1}\mathbf{X}_{12}^{*\prime}\mathbf{V}_{11}^{-1/2}\mathbf{y}_1 + \mathbf{X}_{21}^{**}\boldsymbol{\beta}_1$, which does not depend on the unknown parameter vector $\boldsymbol{\beta}_1$, if, and only if, $\mathbf{X}_{21}^{**} = \mathbf{0}$. Now, let $\mathbf{A} = \mathbf{A}_2\mathbf{A}_1$ and $\mathbf{Q} = \mathbf{I}$. Then, the linear structure $(\mathbf{A}, \mathbf{Q})$ is sufficient for the BLUE of $\boldsymbol{\beta}_1$ in the sense of Gourieroux and Monfort (1980) if, and only if, $\mathbf{X}_{21}^{**} = \mathbf{0}$, which, by the definition of $\mathbf{X}_{21}^{**}$ in (14), is just the necessary and sufficient condition given in Theorem 3. This conditional distribution interpretation of the partially superfluous condition (12) of Theorem 3 is in principle the same as the approach used by Gourieroux and Monfort (1980, p. 1093) for deriving their necessary and sufficient condition for the second subsample to be fully superfluous for the BLUE of the whole parameter vector $\boldsymbol{\beta}$.

Remarks.

(1) It is interesting to observe that the necessary and sufficient condition for the second subsample to be partially superfluous for the BLUE of $\boldsymbol{\beta}_1$ does not depend on $\mathbf{V}_{22}$, as long as the covariance matrix $\sigma^2\mathbf{V}$ is positive definite.

(2) Notice that the fully superfluous condition (5) of Gourieroux and Monfort (1980, p. 1093) is $\mathbf{X}_{2\cdot} = \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{1\cdot}$, which implies $\mathbf{X}_{21} = \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{11}$ and $\mathbf{X}_{22} = \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{12}$. Thus, we can easily see that the fully superflu-

ous condition (5) is sufficient but not necessary for the partially superfluous condition (12).

REFERENCES

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

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

APPENDIX

Proof of Lemma 2. We first note that

$$\begin{aligned}\operatorname{var}(\hat{\boldsymbol\beta}_1)-\operatorname{var}(\tilde{\boldsymbol\beta}_1) &= \sigma^2\mathbf{S}(\mathbf{X}_{1\bullet}'\mathbf{V}_{11}^{-1}\mathbf{X}_{1\bullet})^{-1}\mathbf{S}'-\sigma^2\mathbf{S}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{S}'\\ &= \sigma^2(\mathbf{S},\mathbf{S})\begin{pmatrix}-\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{0}\\ \mathbf{0} & \mathbf{X}_{1\bullet}'\mathbf{V}_{11}^{-1}\mathbf{X}_{1\bullet}\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{S}'\\ \mathbf{S}'\end{pmatrix}.\end{aligned}$$

Thus using $\operatorname{rk}(\mathbf{A})=\operatorname{rk}(-\mathbf{A})$ and Lemma 1(i), we have

$$\begin{aligned}&\operatorname{rk}[\operatorname{var}(\hat{\boldsymbol\beta}_1)-\operatorname{var}(\tilde{\boldsymbol\beta}_1)]\\ &\quad=\operatorname{rk}\left(-(\mathbf{S},\mathbf{S})\begin{pmatrix}-\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{0}\\ \mathbf{0} & \mathbf{X}_{1\bullet}'\mathbf{V}_{11}^{-1}\mathbf{X}_{1\bullet}\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{S}'\\ \mathbf{S}'\end{pmatrix}\right)\\ &\quad=\operatorname{rk}\begin{pmatrix}-\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{0} & \mathbf{S}'\\ \mathbf{0} & \mathbf{X}_{1\bullet}'\mathbf{V}_{11}^{-1}\mathbf{X}_{1\bullet} & \mathbf{S}'\\ \mathbf{S} & \mathbf{S} & \mathbf{0}\end{pmatrix}-\operatorname{rk}(-\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})-\operatorname{rk}(\mathbf{X}_{1\bullet}'\mathbf{V}_{11}^{-1}\mathbf{X}_{1\bullet})\\ &\quad=\operatorname{rk}\left(\begin{pmatrix}\mathbf{0}&\mathbf{0}&\mathbf{S}'\\ \mathbf{0}&\mathbf{0}&\mathbf{S}'\\ \mathbf{S}&\mathbf{S}&\mathbf{0}\end{pmatrix}-\begin{pmatrix}\mathbf{X}'&\mathbf{0}\\ \mathbf{0}&\mathbf{X}_{1\bullet}'\\ \mathbf{0}&\mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{V}&\mathbf{0}\\ \mathbf{0}&-\mathbf{V}_{11}\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{X}&\mathbf{0}&\mathbf{0}\\ \mathbf{0}&\mathbf{X}_{1\bullet}&\mathbf{0}\end{pmatrix}\right)\\ &\qquad-\operatorname{rk}(\mathbf{X})-\operatorname{rk}(\mathbf{X}_{1\bullet})\\ &\quad=\operatorname{rk}\begin{pmatrix}\mathbf{V}&\mathbf{0}&\mathbf{X}&\mathbf{0}&\mathbf{0}\\ \mathbf{0}&-\mathbf{V}_{11}&\mathbf{0}&\mathbf{X}_{1\bullet}&\mathbf{0}\\ \mathbf{X}'&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{S}'\\ \mathbf{0}&\mathbf{X}_{1\bullet}'&\mathbf{0}&\mathbf{0}&\mathbf{S}'\\ \mathbf{0}&\mathbf{0}&\mathbf{S}&\mathbf{S}&\mathbf{0}\end{pmatrix}-\operatorname{rk}(\mathbf{V})-\operatorname{rk}(\mathbf{V}_{11})-2K, \qquad \textbf{(A.1)}\end{aligned}$$

using $\operatorname{rk}(\mathbf{X})=\operatorname{rk}(\mathbf{X}_{1\bullet})=K$. Now, substituting the definition of $\mathbf{S}=(\mathbf{I}_{K_1},\mathbf{0})$ into the block matrix in the first term of (A.1) and simplifying it by elementary block matrix operations, we have

https://doi.org/10.1017/S0266466606060269 Published online by Cambridge University Press

$$
\operatorname{rk}\begin{pmatrix}
\mathbf{V}_{11} & \mathbf{V}_{12} & \mathbf{0} & \mathbf{X}_{11} & \mathbf{X}_{12} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\
\mathbf{V}_{21} & \mathbf{V}_{22} & \mathbf{0} & \mathbf{X}_{21} & \mathbf{X}_{22} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\
\mathbf{0} & \mathbf{0} & -\mathbf{V}_{11} & \mathbf{0} & \mathbf{0} & \mathbf{X}_{11} & \mathbf{X}_{12} & \mathbf{0}\\
\mathbf{X}'_{11} & \mathbf{X}'_{21} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{K_1}\\
\mathbf{X}'_{12} & \mathbf{X}'_{22} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\
\mathbf{0} & \mathbf{0} & \mathbf{X}'_{11} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{K_1}\\
\mathbf{0} & \mathbf{0} & \mathbf{X}'_{12} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{K_1} & \mathbf{0} & \mathbf{I}_{K_1} & \mathbf{0} & \mathbf{0}
\end{pmatrix}
$$

$$
= \operatorname{rk}\begin{pmatrix}
\mathbf{V}_{11} & \mathbf{V}_{12} & \mathbf{0} & \mathbf{X}_{11} & \mathbf{X}_{12} & \mathbf{0}\\
\mathbf{V}_{21} & \mathbf{V}_{22} & \mathbf{0} & \mathbf{X}_{21} & \mathbf{X}_{22} & \mathbf{0}\\
\mathbf{0} & \mathbf{0} & -\mathbf{V}_{11} & -\mathbf{X}_{11} & -\mathbf{X}_{12} & \mathbf{X}_{12}\\
\mathbf{X}'_{11} & \mathbf{X}'_{21} & -\mathbf{X}'_{11} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\
\mathbf{X}'_{12} & \mathbf{X}'_{22} & -\mathbf{X}'_{12} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\
\mathbf{0} & \mathbf{0} & \mathbf{X}'_{12} & \mathbf{0} & \mathbf{0} & \mathbf{0}
\end{pmatrix} + 2K_1
$$

$$
= \operatorname{rk}\begin{pmatrix}
\mathbf{V}_{11} & \mathbf{V}_{12} & \mathbf{V}_{11} & \mathbf{X}_{11} & \mathbf{X}_{12} & \mathbf{0}\\
\mathbf{V}_{21} & \mathbf{V}_{22} & \mathbf{V}_{21} & \mathbf{X}_{21} & \mathbf{X}_{22} & \mathbf{0}\\
\mathbf{V}_{11} & \mathbf{V}_{12} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{X}_{12}\\
\mathbf{X}'_{11} & \mathbf{X}'_{21} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\
\mathbf{X}'_{12} & \mathbf{X}'_{22} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\
\mathbf{0} & \mathbf{0} & \mathbf{X}'_{12} & \mathbf{0} & \mathbf{0} & \mathbf{0}
\end{pmatrix} + 2K_1
$$

$$
= \operatorname{rk}\begin{pmatrix}
\mathbf{V} & \mathbf{V}_1 & \mathbf{X} & \mathbf{0}\\
\mathbf{V}'_1 & \mathbf{0} & \mathbf{0} & \mathbf{X}_{12}\\
\mathbf{X}' & \mathbf{0} & \mathbf{0} & \mathbf{0}\\
\mathbf{0} & \mathbf{X}'_{12} & \mathbf{0} & \mathbf{0}
\end{pmatrix} + 2K_1 \quad (\mathbf{V}_1 \equiv (\mathbf{V}_{11}, \mathbf{V}_{12})')
$$

$$
= \operatorname{rk}\begin{pmatrix}
\mathbf{V} & \mathbf{0} & \mathbf{V}_1 & \mathbf{X}\\
\mathbf{0} & \mathbf{0} & \mathbf{X}'_{12} & \mathbf{0}\\
\mathbf{V}'_1 & \mathbf{X}_{12} & \mathbf{0} & \mathbf{0}\\
\mathbf{X}' & \mathbf{0} & \mathbf{0} & \mathbf{0}
\end{pmatrix} + 2K_1 \quad \text{(interchanging rows and columns)}
$$

$$
= \operatorname{rk}\begin{pmatrix}
\mathbf{V} & \mathbf{V}_1 & \mathbf{X}\\
\mathbf{0} & \mathbf{X}'_{12} & \mathbf{0}
\end{pmatrix} + r\begin{pmatrix}
\mathbf{V}_1 & \mathbf{X}\\
\mathbf{X}'_{12} & \mathbf{0}
\end{pmatrix} + 2K_1 \quad \text{(by (10))}
$$

https://doi.org/10.1017/S0266466606060269 Published online by Cambridge University Press

$$= \text{rk}(\mathbf{V}) + \text{rk}(\mathbf{X}_{12}) + \text{rk}\begin{pmatrix} \mathbf{V}_1 & \mathbf{X} \\ \mathbf{X}'_{12} & \mathbf{0} \end{pmatrix} + 2K_1$$

$$= \text{rk}(\mathbf{V}) + \text{rk}\begin{pmatrix} \mathbf{V}_1 & \mathbf{X} \\ \mathbf{X}'_{12} & \mathbf{0} \end{pmatrix} + 2K - \text{rk}(\mathbf{X}_{12}), \quad \textbf{(A.2)}$$

using $\text{rk}(\mathbf{X}_{12}) = K_2$ and $K = K_1 + K_2$. Now, substituting (A.2) into the first term of (A.1), we easily obtain

$$\text{rk}[\text{var}(\hat{\boldsymbol{\beta}}_1) - \text{var}(\tilde{\boldsymbol{\beta}}_1)] = \text{rk}\begin{pmatrix} \mathbf{V}_1 & \mathbf{X} \\ \mathbf{X}'_{12} & \mathbf{0} \end{pmatrix} - \text{rk}(\mathbf{X}_{12}) - \text{rk}(\mathbf{V}_{11}). \quad \textbf{(A.3)}$$

We now proceed to find the rank of the matrix of the first term on the right-hand side of (A.3). In fact, using the definition of $\mathbf{X}$ and $\mathbf{V}_1 = (\mathbf{V}_{11}, \mathbf{V}_{12})'$ and the formula (9), we have

$$\text{rk}\begin{pmatrix} \mathbf{V}_1 & \mathbf{X} \\ \mathbf{X}'_{12} & \mathbf{0} \end{pmatrix} = \text{rk}\begin{pmatrix} \mathbf{V}_{11} & \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{V}_{21} & \mathbf{X}_{21} & \mathbf{X}_{22} \\ \mathbf{X}'_{12} & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

$$= \text{rk}\begin{pmatrix} \mathbf{V}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_{21} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{11} & \mathbf{X}_{22} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{12} \\ \mathbf{0} & -\mathbf{X}'_{12}\mathbf{V}_{11}^{-1}\mathbf{X}_{11} & -\mathbf{X}'_{12}\mathbf{V}_{11}^{-1}\mathbf{X}_{12} \end{pmatrix}$$

$$= \text{rk}(\mathbf{V}_{11}) + \text{rk}\begin{pmatrix} -\mathbf{X}'_{12}\mathbf{V}_{11}^{-1}\mathbf{X}_{12} & -\mathbf{X}'_{12}\mathbf{V}_{11}^{-1}\mathbf{X}_{11} \\ \mathbf{X}_{22} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{12} & \mathbf{X}_{21} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{11} \end{pmatrix}$$

$$= \text{rk}(\mathbf{V}_{11}) + \text{rk}(-\mathbf{X}'_{12}\mathbf{V}_{11}^{-1}\mathbf{X}_{12})$$
$$+ \text{rk}[(\mathbf{X}_{21} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{11})$$
$$- (\mathbf{X}_{22} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{12})(\mathbf{X}'_{12}\mathbf{V}_{11}^{-1}\mathbf{X}_{12})^{-1}\mathbf{X}'_{12}\mathbf{V}_{11}^{-1}\mathbf{X}_{11}]$$

$$= \text{rk}(\mathbf{V}_{11}) + \text{rk}(\mathbf{X}_{12})$$
$$+ \text{rk}[(\mathbf{X}_{21} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{11})$$
$$- (\mathbf{X}_{22} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{X}_{12})(\mathbf{X}'_{12}\mathbf{V}_{11}^{-1}\mathbf{X}_{12})^{-1}\mathbf{X}'_{12}\mathbf{V}_{11}^{-1}\mathbf{X}_{11}].$$

By substituting it into (A.3), the rank expression given in (11) of Lemma 2 follows immediately. ■