Journal of Multivariate Analysis **70**, 86–94 (1999)
Article ID jmva.1999.1817, available online at http://www.idealibrary.com on IDEAL®

On the Efficiencies of Several Generalized Least Squares Estimators in a Seemingly Unrelated Regression Model and a Heteroscedastic Model

Hiroshi Kurata*

Yamaguchi University, Yamaguchi, Japan
E-mail: kurata@po.cc.yamaguchi-u.ac.jp

Received September 18, 1997

This paper investigates the efficiencies of several generalized least squares estimators (GLSEs) in terms of the covariance matrix. Two models are analyzed: a seemingly unrelated regression model and a heteroscedastic model. In both models, we define a class of unbiased GLSEs and show that their covariance matrices remain the same even if the distribution of the error term deviates from the normal distributions. The results are applied to the problem of evaluating the lower and upper bounds for the covariance matrices of the GLSEs. © 1999 Academic Press

AMS 1991 subject classifications: 62J05.
Key words and phrases: elliptically symmetric distributions; covariance matrix; generalized least squares estimators; seemingly unrelated regression model; heteroscedastic model.

1. INTRODUCTION

In this paper, we consider the linear regression model

$$y = X\beta + \varepsilon \qquad \text{with} \quad E(\varepsilon) = 0 \text{ and } E(\varepsilon\varepsilon') = \Omega, \tag{1}$$

where

$$\Omega = \Omega(\theta) \in \mathcal{S}^{+}(n).$$

Here, X is an $n \times k$ known matrix of full rank, Ω is a function of an unknown vector θ, and $\mathcal{S}^{+}(n)$ denotes the set of $n \times n$ positive definite matrices. A generalized least squares estimator (GLSE) of β is defined as

$$\hat{\beta}(\hat{\Omega}) = (X'\hat{\Omega}^{-1}X)^{-1}\, X'\hat{\Omega}^{-1}y \qquad \text{with} \quad \hat{\Omega} = \Omega(\hat{\theta}), \tag{2}$$

* The author thanks the editors and an anonymous referee for their valuable comments on the earlier version of this paper. This research is financially supported by *Faculty of Economics Research Fund in Yamaguchi University*.

0047-259X/99 $30.00

where $\hat{\theta}$ is an estimator of θ. This paper investigates the finite sample efficiencies of several unbiased GLSEs in terms of covariance matrix. We assume that the distribution of the error term ε is elliptically symmetric with mean 0 and covariance Ω. That is, the probability density function $f(\varepsilon)$ of ε is expressed as

$$f(\varepsilon) = |\Omega|^{-1/2} f_0(\varepsilon'\Omega^{-1}\varepsilon)$$

for some nonnegative function f_0 such that $\int_{R^n} f_0(x'x)\,dx = 1$ and $\int_{R^n} xx' f_0(x'x)\,dx = I_n$. This will be written as $\mathscr{L}(\varepsilon) \in E_n(0, \Omega)$, where $\mathscr{L}(\varepsilon)$ denotes the distribution of ε. The class $E_n(0, \Omega)$ contains some distributions whose tails are longer than that of the normal distribution $N_n(0, \Omega)$.

In the case where ε is distributed as $N_n(0, \Omega)$, several GLSEs satisfy the following condition: conditional on $\hat{\Omega}$,

$$E(\hat{\beta}(\hat{\Omega}) \mid \hat{\Omega}) = \beta \qquad \text{and} \qquad Cov(\hat{\beta}(\hat{\Omega}) \mid \hat{\Omega}) = H, \tag{3}$$

where H is given mby

$$H \equiv H(\hat{\Omega}, \Omega) = (X'\hat{\Omega}^{-1}X)^{-1} X'\hat{\Omega}^{-1}\Omega\hat{\Omega}^{-1}X(X'\hat{\Omega}^{-1}X)^{-1}. \tag{4}$$

Clearly (3) implies that $Cov(\hat{\beta}(\hat{\Omega})) = E(H(\hat{\Omega}, \Omega))$. Typical examples are the unrestricted Zellner estimator (UZE) in a seemingly unrelated regression (SUR) model (Zellner, 1962, 1963) and a GLSE in a heteroscedastic model. For such a GLSE, Kurata and Kariya (1996) and Kariya (1981) derived the lower and upper bounds for the covariance matrix $Cov(\hat{\beta}(\hat{\Omega}))$ as

$$(X'\Omega^{-1}X)^{-1} \leqslant Cov(\hat{\beta}(\hat{\Omega})) \leqslant E[L(\hat{\Omega}, \Omega)](X'\Omega^{-1}X)^{-1}, \tag{5}$$

where $L(\hat{\Omega}, \Omega) = (l_1 + l_n)^2/4l_1 l_n$ and $l_1 \leqslant \cdots \leqslant l_n$ are the latent roots of $\Omega^{-1/2}\hat{\Omega}\Omega^{-1/2}$. Bilodeau (1990) regarded $L(\hat{\Omega}, \Omega)$ as a loss function for choosing an estimator $\hat{\Omega}$ in $\hat{\beta}(\hat{\Omega})$ and derived the optimal estimator with respect to $L(\hat{\Omega}, \Omega)$.

However, it strongly depends on the normality of ε whether a GLSE satisfies the condition (3) or not. Hence we need to investigate the cases where the distribution of ε may deviate from the normality. In the next section, we consider the SUR model under the elliptical symmetry of ε. We first define a class of unbiased GLSE and show that their covariance matrices remain the same as long as $\mathscr{L}(\varepsilon) \in E_n(0, \Omega)$. This leads to a generalization of the results of Kariya (1981), Bilodeau (1990), and Kurata and Kariya (1996) stated above. In Section 3, we pursue a similar analysis for some GLSE in the heteroscedastic model.

In the recent literature, Hasegawa (1994, 1995, 1996) treated the Revankar's SUR model (Revankar, 1974), an SUR model of a simple structure, and investigated the finite sample efficiencies of typical GLSEs under several families of non-normal distributions. See also the references therein. As for the the asymptotic efficiencies of the GLSEs under non-normality, consult Srivastava and Maekawa (1995). Fundamental results of statistical inference in the SUR model are summarized in Srivastava and Giles (1987).

Here we briefly review several facts on the elliptically symmetric distributions. Let an $n \times 1$ random vector x be distributed as $E_n(0, \Omega)$. Then $\mathscr{L}(\Gamma \Omega^{-1/2} x) = \mathscr{L}(\Omega^{-1/2} x)$ holds for any $\Gamma \in \mathcal{O}(n)$, where $\mathcal{O}(n)$ denotes the group of $n \times n$ orthogonal matrices. If $\Omega = I_n$, then $\|x\| = \sqrt{x'x}$ and $x/\|x\|$ are independent and $x/\|x\|$ is distributed as the uniform distribution on the unit sphere in R^n (Muirhead, 1982, Chap. 1). Decompose x as $x = (x_1', x_2')'$ with $x_j : n_j \times 1$ and $n_1 + n_2 = n$. Then the marginal distribution of x_j is n_j-variate elliptical with mean 0 and covariance I_{n_j}. Further, the conditional distribution of x_1 given x_2 is also n_1-variate elliptical with mean

$$E(x_1 \mid x_2) = 0 \tag{6}$$

and conditional covariance

$$Cov(x_1 \mid x_2) = \tilde{c}(x_2' x_2)\, I_{n_1} \tag{7}$$

for some function $\tilde{c}$ (Fang and Zhang, 1990, Chap 2). The function $\tilde{c}$ satisfies $E(\tilde{c}(x_2' x_2)) = 1$, since $Cov(x_1) = I_{n_1}$. Note that if x is normal, then $\tilde{c} \equiv 1$.

2. EFFICIENCIES OF GLSEs IN THE SUR MODEL

The SUR model considered here is the model (1) with the structure

$$\begin{aligned} &y = (y_1', ..., y_N')', \qquad X = diag\{X_1, ..., X_N\},\ \beta = (\beta_1', ..., \beta_N')' \\ &\varepsilon = (\varepsilon_1', ..., \varepsilon_N')', \qquad \Omega = \Sigma \otimes I_m \text{ and } \Sigma \in \mathscr{S}^+(N), \end{aligned} \tag{8}$$

here $y_j : m \times 1$, $X_j : m \times k_j$, $\beta_j : k_j \times 1$, $k = \sum_{j=1}^N k_j$, $n = Nm$ and *diag* denotes the block diagonal matrix. $\mathscr{L}(\varepsilon) \in E_n(0, \Sigma \otimes I_m)$ is assumed.

Let $\hat{\Sigma} \equiv \hat{\Sigma}(S)$ be an estimator of Σ which depends on y only through the random matrix S, where

$$\begin{aligned} S &= Y'[I - X_*(X_*' X_*)^+ X_*']\, Y \\ &= E'[I - X_*(X_*' X_*)^+ X_*']\, E : N \times N, \end{aligned}$$

$Y = (y_1, ..., y_N) : m \times N$, $X_* = (X_1, ..., X_N) : m \times k$, $E = (\varepsilon_1, ..., \varepsilon_N) : m \times N$, and A^+ denotes the Moore–Penrose generalized inverse of A. Let $\widetilde{\mathscr{C}}$ be the class of GLSEs of the form $\hat{\beta}(\hat{\Sigma} \otimes I_m)$ with $\hat{\Sigma} = \hat{\Sigma}(S)$ and $\hat{\Sigma} \in \mathscr{S}^+(N)$ a.s.. Any GLSE in $\widetilde{\mathscr{C}}$ is unbiased if the expectation exists, because S is an even function of the ordinary least squares residual vector (Kariya and Toyooka, 1985, and Eaton, 1985). See also Theorem 2.1 below. We define a subclass $\mathscr{C}$ of $\widetilde{\mathscr{C}}$ as

$$\mathscr{C} = \{ \hat{\beta}(\hat{\Sigma} \otimes I_m) \in \widetilde{\mathscr{C}} \mid \hat{\Sigma}(S) \text{ is one to one continuous function of } S. \text{ For any } a > 0, \text{ there exists a positive number } \gamma \equiv \gamma(a) \text{ such that } \hat{\Sigma}(aS) = \gamma(a)\, \hat{\Sigma}(S) \}.$$

The class $\mathscr{C}$ contains the GLSEs with such $\hat{\Sigma}(S)$'s as

$$\hat{\Sigma}(S) = TDT', \tag{9}$$

where T is the lower triangular matrix with positive diagonal elements such that $S = TT'$ and D is a diagonal matrix with positive elements. In this case, $\gamma(a) = a$. (It seems to be difficult to find reasonable estimators with $\gamma(a) \neq a$.) By letting $D = I_N$ in (9), we can see that the GLSE with $\hat{\Sigma}(S) = S$, the unrestricted Zellner estimator (UZE), is in $\mathscr{C}$. It is easily shown from the general result of Kariya and Toyooka (1985) that any GLSE in $\mathscr{C}$ has finite second moments.

We introduce some notations. Let $p = rank X_*$ and $q = m - p$. Let $\widetilde{X}$ and $\widetilde{Z}$ be any $m \times p$ and $m \times q$ matrices such that

$$\widetilde{X}\widetilde{X}' = X_*(X'_* X_*)^+ X'_*, \quad \widetilde{X}'\widetilde{X} = I_p$$

and

$$\widetilde{Z}\widetilde{Z}' = I_m - X_*(X'_* X_*)^+ X'_*, \qquad \widetilde{Z}'\widetilde{Z} = I_q.$$

Then $\Gamma \equiv (\widetilde{X}, \widetilde{Z}) \in \mathscr{O}(m)$. Let

$$\tilde{\varepsilon} = (\Sigma^{-1/2} \otimes I_m)\, \varepsilon \equiv (\tilde{\varepsilon}'_1, ..., \tilde{\varepsilon}'_N)'$$

$$\eta = (I_N \otimes \Gamma')\, \tilde{\varepsilon} \equiv (\eta'_1, ..., \eta'_N)' \qquad \text{with} \quad \eta_j = \begin{pmatrix} \widetilde{X}'\tilde{\varepsilon}_j \\ \widetilde{Z}'\tilde{\varepsilon}_j \end{pmatrix} = \begin{pmatrix} \delta_j \\ \xi_j \end{pmatrix},$$

where $\tilde{\varepsilon}_j : m \times 1$, $\delta_j : p \times 1$ and $\xi_j : q \times 1$. Then we can easily see that

$$\mathscr{L}(\tilde{\varepsilon}) = \mathscr{L}(\eta) = \mathscr{L}((\delta', \xi')') \in E_n(0, I_n)$$

$$S = \Sigma^{1/2} U' U \Sigma^{1/2},$$

where $\delta = (\delta'_1, ..., \delta'_N)' : Np \times 1$, $\xi = (\xi'_1, ..., \xi'_N)' : Nq \times 1$ and $U = (\xi_1, ..., \xi_N) : q \times N$. As a function of ξ, $S \equiv S(\xi)$ satisfies

$$S(a\xi) = a^2 S(\xi) \qquad \text{for any} \quad a > 0. \tag{10}$$

If ε is normally distributed, then S is distributed as $W_N(\Sigma, q)$, the Wishart distribution with mean $q\Sigma$ and degrees of freedom q.

THEOREM 2.1. *Suppose that* $\mathscr{L}(\varepsilon) \in E_n(0, \Sigma \otimes I_m)$.

(i) *If* $\hat{\beta}(\hat{\Sigma} \otimes I_m) \in \widetilde{\mathscr{C}}$, *then* $E(\hat{\beta}(\hat{\Sigma} \otimes I_m) \mid \hat{\Sigma}) = \beta$.

(ii) *If* $\hat{\beta}(\hat{\Sigma} \otimes I_m) \in \widetilde{\mathscr{C}}$, *then*

$$Cov(\hat{\beta}(\hat{\Sigma} \otimes I_m) \mid \hat{\Sigma}) = c(\hat{\Sigma})\, H(\hat{\Sigma} \otimes I_m, \Sigma \otimes I_m) \tag{11}$$

for some function c *such that* $E(c(\hat{\Sigma})) = 1$.

(iii) *If* $\hat{\beta}(\hat{\Sigma} \otimes I_m) \in \mathscr{C}$, *then* $c(\hat{\Sigma})$ *and* $H(\hat{\Sigma} \otimes I_m, \Sigma \otimes I_m)$ *are independent.*

Proof. By using $X'_j \widetilde{Z} = 0$ and (6), the following two equalities are proved:

$$X'(\hat{\Sigma}^{-1} \otimes I_m)\, \varepsilon = X'(\hat{\Sigma}^{-1} \Sigma^{1/2} \otimes \widetilde{X})\, \delta \tag{12}$$

and

$$E(\delta \mid \hat{\Sigma}) = E[\, E(E(\delta \mid \xi) \mid S) \mid \hat{\Sigma}] = 0. \tag{13}$$

Hence we obtain $E(\hat{\beta}(\hat{\Sigma} \otimes I_m) \mid \hat{\Sigma}) = (X'(\hat{\Sigma}^{-1} \otimes I_m)\, X)^{-1}\, X'(\hat{\Sigma}^{-1} \Sigma^{1/2} \otimes \widetilde{X})$ $E(\delta \mid \hat{\Sigma}) + \beta = \beta$, proving (i). Similarly, $E(\delta\delta' \mid \hat{\Sigma}) = E[\, E(E(\delta\delta' \mid \xi) \mid S) \mid \hat{\Sigma}]$ $= E[\, E(\tilde{c}(\|\xi\|^2) \mid S) \mid \hat{\Sigma}]\, I_{N_p}$ holds for some function $\tilde{c}$ (see (7)). Noting that $\|\xi\|^2 = tr(S\Sigma^{-1})$ and letting $c(\hat{\Sigma}) = E(\tilde{c}(tr(S\Sigma^{-1})) \mid \hat{\Sigma})$ yield

$$E(\delta\delta' \mid \hat{\Sigma}) = c(\hat{\Sigma})\, I_{N_p}. \tag{14}$$

The function c satisfies $E(c(\hat{\Sigma})) = 1$, since $Cov(\delta) = I_{N_p}$. Therefore, from (12) and (14), we obtain

$$\begin{aligned}
&Cov(\hat{\beta} \pm (\hat{\Sigma} \otimes I_m) \mid \hat{\Sigma}) \\
&\quad = (X'(\hat{\Sigma}^{-1} \otimes I_m)\, X)^{-1}\, X'(\hat{\Sigma}^{-1} \Sigma^{1/2} \otimes \widetilde{X}) \\
&\qquad \times\, E(\delta\delta' \mid \hat{\Sigma})(\Sigma^{1/2} \hat{\Sigma}^{-1} \otimes \widetilde{X}')\, X(X'(\hat{\Sigma}^{-1} \otimes I_m)\, X)^{-1} \\
&\quad = c(\hat{\Sigma})(X'(\hat{\Sigma}^{-1} \otimes I_m)\, X)^{-1}\, X'(\hat{\Sigma}^{-1} \Sigma \hat{\Sigma}^{-1} \otimes \widetilde{X}\widetilde{X}') \\
&\qquad \times X(X'(\hat{\Sigma}^{-1} \otimes I_m)\, X)^{-1}.
\end{aligned}$$

Here it is easily proved by a direct calculation that $X'(\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1} \otimes \tilde{X}\tilde{X}')\,X = X'(\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1} \otimes I)\,X$. Thus we establish (ii). To prove (iii), we assume that $\hat{\beta}(\hat{\Sigma} \otimes I_m) \in \mathscr{C}$. As a function of ξ,

$$H(\hat{\Sigma}(S(\xi)) \otimes I_m, \Sigma \otimes I_m) \equiv \tilde{H}(\xi) \tag{15}$$

depends on ξ only through $\xi/\|\xi\|$, since for any $a > 0$,

$$\begin{aligned}\tilde{H}(a\xi) &= H(\hat{\Sigma}(S(a\xi)) \otimes I_m, \Sigma \otimes I_m)\\ &= H(\hat{\Sigma}(a^2 S(\xi)) \otimes I_m, \Sigma \otimes I_m)\\ &= H(\gamma(a^2)\,\hat{\Sigma}(S(\xi)) \otimes I_m, \Sigma \otimes I_m)\\ &= H(\hat{\Sigma}(S(\xi)) \otimes I_m, \Sigma \otimes I_m)\\ &= \tilde{H}(\xi),\end{aligned}$$

where the second equality follows from (10), the third follows from the definition of $\mathscr{C}$, and the fourth follows because $H(a\hat{\Omega}, \Omega) = H(\hat{\Omega}, \Omega)$ holds for any $a > 0$ in general (see (4)). On the other hand, c is a function of $\|\xi\|^2$, because

$$\begin{aligned}c(\hat{\Sigma}) &= E[\tilde{c}(tr(S\Sigma^{-1})) \mid \hat{\Sigma}]\\ &= E[\tilde{c}(tr(S\Sigma^{-1})) \mid S] \qquad (\hat{\Sigma} \text{ is a one to one function of } S)\\ &= \tilde{c}(tr(S\Sigma^{-1})) = \tilde{c}(\|\xi\|^2).\end{aligned}$$

This proves (iii). ∎

Since the distribution of $\xi/\|\xi\|$ is unique, the following theorem is obtained.

THEOREM 2.2. *For any GLSE* $\hat{\beta}(\hat{\Sigma} \otimes I_m) \in \mathscr{C}$, *the covariance matrix* $Cov(\hat{\beta}(\hat{\Sigma} \otimes I_m))$ *remains the same as long as* $\mathscr{L}(\varepsilon) \in E_n(0, \Sigma \otimes I_m)$.

It follows from this result and (5) that any GLSE $\hat{\beta}(\hat{\Sigma} \otimes I_m) \in \mathscr{C}$ satisfies

$$(X'(\Sigma^{-1} \otimes I_m)\,X)^{-1} \leqslant Cov(\hat{\beta}(\hat{\Sigma} \otimes I_m)) \leqslant E[L(\hat{\Sigma}, \Sigma)](X'(\Sigma^{-1} \otimes I_m)\,X)^{-1}$$
$$\text{with} \quad L(\hat{\Sigma}, \Sigma) = (l_1 + l_N)^2/4 l_1 l_N, \tag{16}$$

where $l_1 \leqslant \cdots \leqslant l_N$ are the latent roots of $\Sigma^{-1/2}\hat{\Sigma}(S)\,\Sigma^{-1/2}$ and the expectation $E[L(\hat{\Sigma}, \Sigma)]$ is calculated under $S \sim W_N(\Sigma, q)$. In particular, when $N = 2$, the upper bound for the covariance matrix of the UZE is given by $1 + 2/(q-3)$ (Kariya, 1981). For general N, see Kurata and Kariya (1996). As a loss function for estimating Σ, $L(\hat{\Sigma}, \Sigma)$ is invariant under the group $\mathscr{G}\mathscr{T}^+(N)$ of $N \times N$ lower triangular matrices with positive diagonal elements with action $\hat{\Sigma} \to G\hat{\Sigma}G'$ and $\Sigma \to G\Sigma G'$, where $G \in \mathscr{G}\mathscr{T}^+(N)$. An equivariant estimator of Σ is of the form (9). In the case where $N=2$, Bilodeau (1990) derived the optimal estimator of Σ with respect to this loss function as

$$\hat{\Sigma}_B \equiv \hat{\Sigma}_B(S) = TD_BT' \qquad \text{with} \quad D_B = diag\{1, \sqrt{(q+3)/(q-1)}\}$$

and proposed the GLSE $\hat{\beta}(\hat{\Sigma}_B \otimes I_m)$. Since $L(\hat{\Sigma}(S(a\xi)), \Sigma) = L(\gamma(a^2)\, \hat{\Sigma}(S(\xi)), \Sigma) = L(\hat{\Sigma}(S(\xi)), \Sigma)$ holds for any $a>0$, the upper bounds themselves also remain the same as long as $\mathscr{L}(\varepsilon) \in E_n(0, \Sigma \otimes I_m)$. This implies that the GLSE $\hat{\beta}(\hat{\Sigma}_B \otimes I_m)$ is still optimal under the elliptical symmetry of ε.

3. EFFICIENCIES OF GLSEs IN THE HETEROSCEDASTIC MODEL

In this section, we consider the heteroscedastic model of the form

$$\begin{gathered} y = (y_1', ..., y_N')', \qquad X = (X_1', ..., X_N')', \qquad \varepsilon = (\varepsilon_1', ..., \varepsilon_N')' \\ \Omega = \Omega(\theta) = diag\{\theta_1 I_{m_1}, ..., \theta_N I_{m_N}\} \qquad \text{with} \quad \theta = (\theta_1, ..., \theta_N)' : N \times 1, \end{gathered} \tag{17}$$

where $y_j : m_j \times 1$, $X_j : m_j \times k$, $\varepsilon_j : m_j \times 1$, $n = \sum_{j=1}^N m_j$, and $\mathscr{L}(\varepsilon) \in E_n(0, \Omega(\theta))$. Since the analysis in this section is quite similar to that of Section 2, we often omit the details.

Let $\hat{\theta} = \hat{\theta}(s)$ be an estimator of θ which depends on y only through the random vector $s = (s_1, ..., s_N)' : N \times 1$, where

$$s_j = y_j' \, [I_{m_j} - X_j(X_j' X_j)^+ X_j'] \, y_j / q_j, \tag{18}$$

$p_j = rank X_j$, and $q_j = m_j - p_j$. Let $\tilde{\mathscr{C}}$ be the class of unbiased GLSEs of the form $\hat{\beta}(\Omega(\hat{\theta}))$ with $\Omega(\hat{\theta}) \in \mathscr{S}^+(n)$ a.s.. We consider the following subclass $\mathscr{C}$ of $\tilde{\mathscr{C}}$ whose definition is similar to that of Section 2:

$$\begin{aligned} \mathscr{C} = \{\hat{\beta}(\Omega(\hat{\theta})) \in \tilde{\mathscr{C}} \mid \hat{\theta}(s) &\text{ is one to one continuous function of } s. \text{ For any} \\ & a>0, \text{ there exists a positive number } \gamma \equiv \gamma(a) \text{ such that} \\ & \hat{\theta}(as) = \gamma(a)\, \hat{\theta}(s)\}. \end{aligned}$$

This class contains the GLSEs with such $\hat{\theta}(s)$'s as

$$\hat{\theta}(s) = (f_1 s_1, ..., f_N s_N)', \tag{19}$$

where f_j's are positive constants.

Let $\tilde{X}_j$ and $\tilde{Z}_j$ be any $m_j \times p_j$ and $m_j \times q_j$ matrices such that

$$\tilde{X}_j \tilde{X}_j' = X_j(X_j' X_j)^+ X_j', \qquad \tilde{X}_j' \tilde{X}_j = I_{p_j}$$

and

$$\tilde{Z}_j \tilde{Z}_j' = I_{m_j} - X_j (X_j' X_j)^+ X_j', \qquad \tilde{Z}_j' \tilde{Z}_j = I_{q_j}.$$

Then $\Gamma \equiv diag\{\Gamma_1, ..., \Gamma_N\} \in \mathcal{O}(n)$, where $\Gamma_j = (\tilde{X}_j, \tilde{Z}_j) \in \mathcal{O}(m_j)$. The following notations are essentially the same as those of Section 2,

$$\tilde{\varepsilon} = \Omega^{-1/2} \varepsilon \equiv (\tilde{\varepsilon}_1', ..., \tilde{\varepsilon}_N')'$$

$$\eta = \Gamma' \tilde{\varepsilon} \equiv (\eta_1', ..., \eta_N')' \qquad \text{with} \quad \eta_j = \begin{pmatrix} \tilde{X}_j' \tilde{\varepsilon}_j \\ \tilde{Z}_j' \tilde{\varepsilon}_j \end{pmatrix} \equiv \begin{pmatrix} \delta_j \\ \xi_j \end{pmatrix},$$

where $\tilde{\varepsilon}_j : m_j \times 1$, $\delta_j : p_j \times 1$ and $\xi_j : q_j \times 1$. Let $\delta = (\delta_1', ..., \delta_N')' : (p_1 + \cdots + p_N) \times 1$, $\xi = (\xi_1', ..., \xi_N')' : (q_1 + \cdots + q_N) \times 1$. As a function of ξ, $s \equiv s(\xi)$ satisfies $s(a\xi) = a^2 s(\xi)$ for any $a > 0$.

THEOREM 3.1. *Suppose that* $\mathcal{L}(\varepsilon) \in E_n(0, \Omega(\theta))$.

(i) *If* $\hat{\beta}(\Omega(\hat{\theta})) \in \tilde{\mathscr{C}}$, *the* $E(\hat{\beta}(\Omega(\hat{\theta})) \mid \hat{\theta}) = \beta$.

(ii) *If* $\hat{\beta}(\Omega(\hat{\theta})) \in \tilde{\mathscr{C}}$, *then*

$$Cov(\hat{\beta}(\Omega(\hat{\theta})) \mid \hat{\theta}) = c(\hat{\theta})\, H(\Omega(\hat{\theta}), \Omega(\theta)) \tag{20}$$

for some function c such that $E(c(\hat{\theta})) = 1$.

(iii) *If* $\hat{\beta}(\Omega(\hat{\theta})) \in \mathscr{C}$, *then* $c(\hat{\theta})$ *and* $H(\Omega(\hat{\theta}), \Omega(\theta))$ *are independent.*

Proof. Let $\tilde{X} = diag\{\tilde{X}_1, ..., \tilde{X}_N\}$. Then we can see that $X' \Omega(\hat{\theta})^{-1} \varepsilon = X' \Omega(\hat{\theta})^{-1} \Omega(\theta)^{1/2} \tilde{X} \delta$. This proves (i). The proofs of (ii) and (iii) are parallel to those of Theorem 2.1. ∎

THEOREM 3.2. *For any GLSE* $\hat{\beta}(\Omega(\hat{\theta})) \in \mathscr{C}$, *the covariance matrix* $Cov(\hat{\beta}(\Omega(\hat{\theta}))$ *remains the same as long as* $\mathcal{L}(\varepsilon) \in E_n(0, \Omega(\theta))$.

The implications of this theorem are also similar to those of Theorem 2.2 and omitted.

REFERENCES

1. M. Bilodeau, On the choice of the loss fuction in covariance estimation, *Statist. Decisions* **8** (1990), 131–139.
2. M. L. Eaton, The Gauss–Markov theorem in multivariate analysis, *in* "Multivariate Analysis" (P. R. Krishnaiah, Ed.), Vol. 6, pp. 177–201, 1985.
3. K. T. Fang and Y. T. Zhang, "Generalized Multivariate Analysis," Springer-Verlag, New York/Berlin, 1990.

4. H. Hasegawa, On small sample properties of estimators after a preliminary test of independence in a constrained bivariate linear regression model, *Econom. Stud. Quart.* **45** (1994), 306–320.
5. H. Hasegawa, On small sample properties of Zellner's estimator for the case of two SUR equations with compound normal disturbances, *Comm. Statist. Simulation Comput.* **24** (1995), 45–59.
6. H. Hasegawa, On small sample properties of Zellner's estimator for two SUR equations model with specification error, *J. Japan. Statist. Soc.* **26** (1996), 25–40.
7. T. Kariya, Bounds for the covariance matrices of Zellner's estimator in the SUR model and 2SAE in a heteroscedasitic model, *J. Amer. Statist. Assoc.* **76** (1981), 975–979.
8. T. Kariya and Y. Toyooka, Nonlinear versions of the Gauss–Markov Theorem and GLSE, *in* "Multivariate Analysis" (P. R. Krishnaiah, Ed.), Vol. 6, pp. 345–354, 1985.
9. H. Kurata and T. Kariya, Least upper bound for the covariance matrix of a generalized least squares estimator in regression with applications to a seemingly unrelated regression model and a heteroscedastic model, *Ann. Statist.* **24** (1996), 1547–1559.
10. R. J. Muirhead, "Aspects of Multivariate Statistical Theory," Wiley, New York, 1982.
11. N. S. Revankar, Some finite sample results in the context of two seemingly unrelated regression equations, *J. Amer. Statist. Assoc.* **69** (1974), 634–639.
12. V. K. Srivastava and D. E. A. Giles, "Seemingly Unrelated Regression Equation Models," Dekker, New York, 1987.
13. V. K. Srivastava and K. Maekawa, Efficiency properties of feasible generalized least squares estimators in SURE models under non-normal disturbances, *J. Econometrics* **66** (1995), 99–121.
14. A. Zellner, An efficient method of estimating seemingly unrelated regression and tests for aggregation bias, *J. Amer. Statist. Assoc.* **57** (1962), 348–368.
15. A. Zellner, Estimators for seemingly unrelated regression equations: Some exact sample results, *J. Amer. Statist. Assoc.* **58** (1963), 977–992.