Ann. Inst. Statist. Math.
34 (1982), Part A, 281-297

A METHOD FOR APPROXIMATIONS TO THE PDF'S AND CDF'S OF GLSE'S AND ITS APPLICATION TO THE SEEMINGLY UNRELATED REGRESSION MODEL*

TAKEAKI KARIYA AND KOICHI MAEKAWA

(Received Mar. 23, 1981; revised Nov. 16, 1981)

This paper first develops a valid method for approximations to the pdf's and cdf's of GLSE in linear models and, applying this method to the Zellner estimator with an unrestricted sample covariance in the seemingly unrelated regression model, obtains an approximate pdf with a bound of order n^{-2} and an approximate covariance matrix with a bound of order n^{-3}.

1. Introduction and summary

The validity of Edgeworth type expansions is being recently focussed upon and some general theorems are being established in mathematical statistics and applied fields. Among others, with a clear view over this field, Bhattacharya and Ghosh [1] gave a neat condition for the validity of a formal Edgeworth expansion in the i.i.d. case. Feller [4] also provided a weak condition for the validity in the i.i.d. case. In non i.i.d. but univariate cases, Sargan [12] and Philips [9] derived conditions for the validity in general settings for econometric applications. Further, Durbin [3] extended Feller's result to a non i.i.d. case and applied it to a time series model. Other references are found in these papers. In this paper, taking a full advantage of a special structure of the model concerned here, we give a bound for approximation with the validity. To state our model, let

$$(1.1) \qquad y=X\beta+u \qquad \text{with } u\sim N(0, \Omega)$$

be a linear normal regression model where X is an $n\times k$ fixed matrix of rank k and $\Omega \in S(n)$. Here $S(n)$ denotes the set of all positive defi-

* This research was done at the London School of Economics while the authors were British Council scholars. Kariya is grateful to Professor J. Durbin for a general discussion on asymptotic expansions. Further the authors deeply appreciate Professor Y. Kataoka and anonymous referee for their invaluable comments and suggestions.

nite matrices. By a GLSE we shall mean an estimator of β of the form

$$\hat{\beta} \equiv \hat{\beta}(\hat{\Omega}) = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}y \tag{1.2}$$

where $\hat{\Omega}$ is an estimator of Ω and $\hat{\Omega}$ may be estimated based on a different sample. When Ω is of certain structure, estimators of this form are often used in practice, for example in the SUR (seemingly unrelated regression) model where $\Omega = \Sigma \otimes I$, or in a heteroscedastic model where Ω is a diagonal matrix, or in a serially correlated model where Ω is a function of variance and correlation. However, our method limits its applications to the case that the following assumption holds:

$$\hat{\beta} \mid \hat{\Omega} \sim N\left(\beta, \frac{1}{n} H\right) \tag{1.3}$$

where

$$H = n(X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}\Omega\hat{\Omega}^{-1}X(X'\hat{\Omega}^{-1}X)^{-1} \tag{1.4}$$

and "$\hat{\beta} \mid \hat{\Omega} \sim$" reads "$\hat{\beta}$ given $\hat{\Omega}$ is distributed as". When $\hat{\Omega}$ is estimated independently of y, (1.3) is always satisfied. Further, as is well known, certain GLSE's in the SUR model and a heteroscedastic model also satisfy (1.3). But the GLSE's in serially correlated models do not satisfy (1.3) except the case that an independent sample for estimation of Ω is available. Based on this model, we consider the problems of approximations to the pdf and moments of

$$d = \sqrt{n}(\hat{\beta} - \beta) = \sqrt{n}(X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}u\,. \tag{1.5}$$

In Section 2, after some remarks on the usual approach to such a problem, a bound for the difference between the pdf, say $f(x)$, of d in (1.5) and its approximation, say $f_0(x)$, is evaluated in the form

$$\sup_x |f(x) - f_0(x)| \leq K/n^r \tag{1.6}$$

where the constant K depends on Ω. In Section 3, approximations to the moments of d are treated. It is noted that throughout this paper, the sample size n (and the sample size, say n', for $\hat{\Omega}$ if Ω is independently estimated) is (are) arbitrarily fixed, and so the results hold for any n (n') although too small sample makes the approximations meaningless. To relate our results to those so far obtained, we sometimes use the notation $O(\cdot)$ to show the orders of approximations, but for example, by $Z_n = O(n^{-r})$ we simply mean that $n^r Z_n$ is bounded as a function of n, or $|n^r Z_n| \leq K$ for all n, and we never mean that n is large.

In Section 4, the above method is applied to the UZE (Zellner estimator with the unrestricted covariance matrix) in the 2 equations SUR model, and the approximate covariance matrix up to $O(n^{-2})$ with a bound of order n^{-3} and an approximate pdf up to $O(n^{-1})$ with a bound of order n^{-2} are obtained. In addition, an approximate cdf of a linear combination of the UZE is also obtained with a bound of order n^{-2}. For asymptotic covariance matrices of the UZE and the RZE (Zellner estimator with the restricted sample covariance matrix), Srivastava [13] gave the asymptotic covariance matrix of the RZE up to $O(n^{-1})$ in terms of in-probability concept (see Section 3 for the exact meaning), and interestingly pointed out the equivalence between the asymptotic covariance matrices up to $O(n^{-1})$ of the RZE and UZE. This equivalence is explicitly confirmed by Hall [5] and Srivastava and Upadhyaya [15]. Consequently, to compare these two estimators in covariance matrices, higher order approximations are necessary. The approximate covariance matrix derived here will serve for this purpose as well as it approximates the exact covariance matrix more closely. A different comparison between the UZE and the RZE has been made by Revankar [11], while Kunitomo [8] has derived the exact covariance matrix of the UZE, which seems analytically intractable for such comparisons. Further, some arguments on approximations to covariance matrices are found in Taylor [16] where a 2SAE in a heteroscedastic model is treated.

For approximations to the cdf of d in the SUR model, Philips [10] derived an approximate cdf (cumulative distribution function) of a linear combination of the elements of the d where the validity is based on Sargan [12]. His result is effective only in a neighbourhood of the origin and seems to need that n is large enough to satisfy the conditions for the validity. On the other hand, not only our approximation to the cdf of d is effective globally but also a uniform bound for the approximation is given as a function of n. Further, in his case only a univariate case can be treated, but in this paper we also give a uniform approximation to the joint pdf of the elements of d. Finally, we remark that the method and analyses developed here are applicable to a 2SAE in a heteroscedastic model.

2. Approximations to the pdf of d in (1.5)

For comparisons, we first remark on the so called δ-method frequently used in asymptotic expansions of pdf's or cdf's. In this method, based on the fact that $p\lim \Delta=0$ where $\Delta=\hat{\Omega}-\Omega$, first $\hat{\Omega}^{-1}=\Omega^{-1}(I+\Delta\Omega^{-1})^{-1}$ and then $(X'\hat{\Omega}^{-1}X)^{-1}$ are expanded as infinite series in terms of in-probability, and for d in (1.5), such a form as

(2.1) $$d=\sum_{i=0}^{r-1} d_i+R \qquad \text{where } d_i=O_p(n^{-i/2}) \text{ and } R=O_p(n^{-r/2})$$

is obtained, where O_p denotes the stochastic order. The remainder R is of the form $R=\sum_{i=r}^{\infty} d_i$. Then the characteristic function of d

(2.2) $$\phi(t)=\mathrm{E}\,[\exp(it'd)]$$

is expanded as an infinite series and simplified in the order of $\mathrm{E}\,(O_p(n^{-i/2}))$ with a remainder term of the form $\mathrm{E}\,[O_p(n^{-r/2}, t)]$, which is a function of n and t. Finally by applying a formal Fourier inversion, an approximate pdf or cdf is obtained up to $O(n^{-(r-1)/2})$. Contrary to such many practices, little is known on the validity of this approach. Difficulties lie in the treatment of the remainder terms. In fact, the remainder R in (2.1) is only defined in terms of in-probability, and it is not defined, for example, on the set $\{ch_{\max}(\Delta\Omega^{-1})>1\}$ since on this set $\sum_{j=0}^{\infty}(\Delta\Omega^{-1})^j$ neither converges nor equals $(I+\Delta\Omega^{-1})^{-1}$, where $ch_{\max}(\Delta\Omega^{-1})$ denotes the maximum latent root of $\Delta\Omega^{-1}$. Another difficulty is to show that the remainder term $\mathrm{E}\,[O_p(n^{-r/2}, t)]$ in the expanded characteristic function exists and equals $O(n^{-r/2}, t)$ and that the Fourier inversion of this term is $O(n^{-r/2}, x)$. Thirdly, even if the validity of this procedure is verified, as far as the derivation is based on such an argument that there exists an n_0 such that for $n\geqq n_0$, the result is effective, the final result thus obtained is not free from the restriction $n\geqq n_0$. Here n_0 is usually unknown and may need to be very large.

In our approach, the concept of in-probability is not used at all and a bound for approximation is provided for each n. It is noted that the argument below does not depend on whether or not $\hat{\Omega}$ is estimated based on y only or an independent sample. Let

(2.3) $$A=X'\Omega^{-1}X/n\,, \qquad \bar{X}=\Omega^{-1/2}XA^{-1/2}/\sqrt{n} \text{ and } P=\Omega^{-1/2}\hat{\Omega}\Omega^{-1/2}$$

where for $C\in S(k)$, $C^{1/2}$ satisfies $(C^{1/2})^2=C$. The $\bar{X}'\bar{X}=I_k$ and from (1.4)

(2.4) $$H=A^{-1/2}(\bar{X}'P^{-1}\bar{X})^{-1}\bar{X}'P^{-2}\bar{X}(\bar{X}'P^{-1}\bar{X})^{-1}A^{-1/2}\,.$$

It is noted that replacing P by αP in (2.4) leaves H invariant for $\alpha>0$. Now from (1.3), $d\,|\,\hat{\Omega}\sim N(0, H)$, which implies $\mathrm{Cov}\,(d)=\mathrm{Cov}\,(\sqrt{n}\,\hat{\beta}(\hat{\Omega}))=\mathrm{E}\,(H)$. The next lemma, which is due to Kariya [7], plays a key role below, and so the proof is given.

LEMMA 2.1. $J\equiv H-A^{-1}\geqq 0$.

PROOF. Since $I_n-\bar{X}\bar{X}'$ is idempotent from $\bar{X}'\bar{X}=I_k$, the result follows from

(2.5) $$J=A^{-1/2}\{(\bar{X}'P^{-1}\bar{X})^{-1}\bar{X}'P^{-1}[I-\bar{X}\bar{X}']P^{-1}\bar{X}(\bar{X}'P^{-1}\bar{X})^{-1}\}A^{-1/2}\,.$$

Let

(2.6) $$\bar{V}=P^{-1}-I\,,\qquad D=\bar{X}'\bar{V}\bar{X}\,,\qquad E=\bar{X}'\bar{V}^2\bar{X}$$
$$G=E-D^2\quad\text{and}\quad Q=(I+D)^{-1}=(\bar{X}'P^{-1}\bar{X})^{-1}\,.$$

Then from $\bar{X}'\bar{X}=I_k$ and (2.4)

(2.7) $$H=A^{-1/2}\{(I+D)^{-1}(I+2D+E)(I+D)^{-1}\}A^{-1/2}\,.$$

LEMMA 2.2. $H=A^{-1}+J=A^{-1}+J_1+R_1$ *where*

(2.8) $$J_1=A^{-1/2}(E-D^2)A^{-1/2}\qquad\textit{and}$$
$$R_1=A^{-1/2}[2D^3-DE-ED+GD^2Q+QD^2G+QDGDQ]A^{-1/2}\,.$$

PROOF. The proof is given in Appendix A.

The following theorems state our main results in this section. The proofs are given later.

THEOREM 2.1. *Let $f(x)$ be the pdf of d in* (1.5) *and assume that* $\mathrm{E}\,(t'Jt)^2$ *exists for all t. Let* $\bar{J}_1=\mathrm{E}\,(J_1)$, $\bar{R}_1=\mathrm{E}\,(R_1)$,

(2.9) $$f_0(x)=[1+(1/2)\,\mathrm{tr}\,A\bar{J}_1-(1/2)x'A\bar{J}_1Ax](2\pi)^{-k/2}|A|^{1/2}\exp\,(-(1/2)x'Ax)\,.$$

Then with $c(k,2)=(2\pi)^{-k}2^{-3}$

(2.10) $$\sup_x|f(x)-f_0(x)|\leqq c(k,2)\int\exp\,(-(1/2)t'A^{-1}t)[\mathrm{E}\,(t'Jt)^2+4|t'\bar{R}_1t|]dt\,.$$

It is noted that the approximation in (2.10) is uniform in x, and that it holds for any n although it may be meaningless for n small. Since J_1 is quadratic in $\bar{V}$ as in Lemma 2.3, if $\mathrm{E}\,(J_1)=O(n^{-1})$ and $\mathrm{E}\,(R_1)=O(n^{-3/2})$, $f_0(x)$ in (2.9) approximates $f(x)$ up to $O(n^{-1})$ uniformly with the bound of $O(n^{-3/2})$. Further if the third moments of $\bar{V}$ vanish so that $\mathrm{E}\,(R_1)=O(n^{-2})$ the bound in (2.10) is $O(n^{-2})$. However it is still difficult to evaluate the exact expectations in the bounds, although it depends on the distribution of $\hat{\Omega}$ and the problem concerned.

Next, we consider approximations to the pdf and cdf of a linear combination of d. Let

(2.11) $$z=a'd=\sqrt{n}\,a'[\hat{\beta}(\hat{\Omega})-\beta]\,,$$

where a is a $k\times 1$ fixed vector ($\neq 0$).

THEOREM 2.2. (i) *Let $g(z)$ be the pdf of z in* (2.11) *and assume that* $\mathrm{E}\,(a'Ja)^2$ *exists. Let* $\tau=a'A^{-1}a$,

$$(2.12)\qquad g_0(z)=[1+(1/2)(\tau^{-1}-z^2\tau^{-2})\,\mathrm{E}\,(a'J_1a)](2\pi\tau)^{-1/2}\exp(-z^2/2\tau)\,.$$

Then with $c(1,2)=(2\pi)^{-1}2^{-3}$

$$(2.13)\qquad \sup_z |g(z)-g_0(z)|\leqq c(1,2)\int \exp(-s^2\tau/2)[s^4\,\mathrm{E}\,(a'Ja)^2 + 4s^4|\mathrm{E}\,(a'R_1a)|]ds\,.$$

(ii) *Further let* $G(z)$ *be the cdf of* z *and let* $G_0(z)=\int_{-\infty}^{z} g_0(x)dx$. *Then*

$$(2.14)\qquad \sup_z |G(z)-G_0(z)|\leqq \pi^{-1}[2^{-1}\tau^{-2}\,\mathrm{E}\,(a'Ja)^2+\tau^{-1}\,\mathrm{E}\,|a'R_1a|]\,.$$

For example, choosing $a=(1,0,\cdots,0)$ in (2.13), an approximate pdf or cdf of the first coefficient $\hat{\beta}_{11}(\hat{\Omega})$ of $\hat{\beta}(\hat{\Omega})$ is obtained.

We note that so far we have assumed neither that $A\equiv X'\Omega^{-1}X/n=O(1)$ nor that $\hat{\Omega}$ does guarantee the consistency of $\hat{\beta}(\hat{\Omega})$. This is because the argument in the above has been made for a fixed n. However, the approximation in (2.10) does not make sense unless the bound in (2.10) goes to 0 as $n\to\infty$ (or $n'\to\infty$), where n' is the sample size for estimation of Ω if $\hat{\Omega}$ is estimated independently. From (2.10), the condition for this is that both $\mathrm{E}\,(t'Jt)^2$ and $\mathrm{E}\,(R_1)$ in (2.10) go to 0 for all t as $n\to\infty$ (or $n'\to\infty$). Since $0\leqq \mathrm{E}\,(t'Jt)=\mathrm{E}\,|t'Jt|\leqq[\mathrm{E}\,(t'Jt)^2]^{1/2}$ and $\mathrm{E}\,(t'Jt)=\mathrm{E}\,(t'J_1t)+\mathrm{E}\,(t'R_1t)$ for all t, these conditions imply $\mathrm{E}\,(J_1)\to 0$ as $n\to\infty$ (or $n'\to\infty$). Hence, from Theorem 2.1, these imply that $d=\sqrt{n}\,[\hat{\beta}(\hat{\Omega})-\beta]\to N(0,\lim_n A^{-1})$ in distribution, where $\lim_n A>0$ is assumed here. Therefore, we obtain

LEMMA 2.3. *A necessary condition for the bound in* (2.10) *goes to* 0 *as* $n\to\infty$ (*or* $n'\to\infty$) *is that the limiting distribution of* $d=\sqrt{n}\,[\hat{\beta}(\hat{\Omega})-\beta]$ *is* $N(0,\lim_n A^{-1})$.

Consequently, $\hat{\Omega}$ must guarantee this. Although even for the OLSE $\hat{\beta}(I)$, the approximation (2.10) holds, the bound does not converge to 0 since $\sqrt{n}\,[\hat{\beta}(I)-\beta]\to N(0,\lim_n A^{-1})$. In usual situations where Ω is of some structure, say $\Omega=\Omega(\theta)$, choosing a consistent estimator of θ will guarantee the condition in Lemma 2.3, and further at least the order of the bound will be directly evaluated.

PROOFS OF THEOREMS 2.1 AND 2.2. Using Lemma 2.1, we shall evaluate the characteristic function $\phi(t)$ of d. From (1.3), (2.2) and Lemma 2.1,

$$(2.15)\quad \phi(t)=\mathrm{E}\,[\exp(-(1/2)t'Ht)]=\exp(-(1/2)t'A^{-1}t)\,\mathrm{E}\,[\exp(-(1/2)'tJt)]\,.$$

Assuming that $\mathrm{E}(t'Jt)^r$ exists, a Taylor expansion of (2.15) yields

$$(2.16)\qquad \phi(t)=\phi_1(t)+\mathrm{E}\,[(-(1/2)t'Jt)^r \exp(-(1/2)\theta t'Jt)/r!]$$

where θ is a random variate such that $0\leqq\theta\leqq 1$ and

$$(2.17)\qquad \phi_1(t)=\exp(-(1/2)t'A^{-1}t)\left[\sum_{j=0}^{r-1}(-2)^{-j}\,\mathrm{E}\,(t'Jt)^j/j!\right].$$

Define

$$(2.18)\qquad f_1(x)=(2\pi)^{-k}\int \exp(it'x)\phi_1(t)dt\,.$$

LEMMA 2.4. *Let $f(x)$ be the pdf of d and assume that $\mathrm{E}(t'Jt)^r$ exists for all t. Then with $c(k, r)=(2\pi)^{-k}2^{-r}(r!)^{-1}$*

$$(2.19)\qquad |f(x)-f_1(x)|\leqq c(k, r)\int \mathrm{E}\,(t'Jt)^r \exp(-(1/2)t'A^{-1}t)dt\,.$$

PROOF. From (2.16), (2.17) and (2.18), $|f(x)-f_1(x)|\leqq(2\pi)^{-k}\int|\phi(t)-\phi_1(t)|dt\leqq c(k, r)\int \exp(-(1/2)t'A^{-1}t)[\mathrm{E}\,(t'Jt)^r\exp(-(1/2)\theta t'Jt)]dt$. Since $J\geqq 0$ from Lemma 2.1 and since $0\leqq\theta\leqq 1$, $\exp(-(1/2)\theta t'Jt)\leqq 1$. Hence the result follows.

Now to prove Theorem 2.1, let $r=2$ in (2.17) and (2.19). Then from (2.17), $\phi_1(t)=\phi_0(t)+\exp(-(1/2)t'A^{-1}t)[-(1/2)t'\bar{R}_1t]$ where $\phi_0(t)=\exp(-(1/2)t'A^{-1}t)[1-(1/2)t'\bar{J}_1t]$. Since $f_0(x)=(2\pi)^{-k}\int\exp(-it'x)\phi_0(t)dt$, in the same way as the proof of Lemma 2.4, $|f(x)-f_0(x)|$ is bounded above by the right side of (2.10), completing the proof of Theorem 2.1. To prove Theorem 2.2 (i) let $\Psi(s)$ be the characteristic function of z and let $\Psi_1(s)=\Psi_0(s)+\exp(-s^2\tau/2)[-(1/2)s^2\,\mathrm{E}\,(a'R_1a)]$ where $\Psi_0(s)=\exp(-s^2\tau/2)\cdot[1-(1/2)s^2\,\mathrm{E}\,(a'J_1a)]$. Since for $\phi(t)$, $\phi_0(t)$ and $\phi_1(t)$ in the proof of Theorem 2.1, $\phi(sa)=\Psi(s)$ and $\phi_i(sa)=\Psi_i(s)$ $(i=0, 1)$ hold, the result follows from the proof of Theorem 2.1. To prove Theorem 2.2 (ii), we use the following lemma.

LEMMA 2.5 (Feller [4], p. 512). *Let G be a cdf with mean 0 and characteristic function Ψ. Suppose that $G-G_0$ vanishes at $\pm\infty$ and G_0 has a derivative g_0 such that $|g_0|\leqq K$. Finally, suppose that g_0 has a continuously differentiable Fourier transform Ψ_0 such that $\Psi_0(0)=1$ and $\Psi_0'(0)=0$. Then for all z and $T>0$*

$$(2.20)\qquad |G(z)-G_0(z)|\leqq\frac{1}{\pi}\int_{-T}^{T}\left|\frac{\Psi(s)-\Psi_0(s)}{s}\right|ds+\frac{24K}{T}\,.$$

PROOF OF THEOREM 2.2 (ii). Define $\Psi(s)$, $\Psi_1(s)$ and $\Psi_0(s)$ as in the proof of Theorem 2.2 (i). Then the conditions for $G_0(z)$ and $\Psi_0(s)$ in Lemma 2.5 are easily seen to be satisfied. Hence the left-hand side of (2.20) is bounded above by $\pi^{-1}\int_{-\infty}^{\infty}|(\Psi(s)-\Psi_0(s))/s|ds$. But since $\Psi(s)=\phi(sa)$ and $\Psi_0(s)=\phi_0(sa)$, from (2.16) we obtain

$$|\Psi(s)-\Psi_0(s)|\leqq\exp(-s^2\tau/2)[2^{-3}s^4\,\mathrm{E}\,(a'Ja)^2+2^{-1}s^2\,\mathrm{E}\,|a'R_1a|]\ .$$

Using this inequality and the fact that when $s\sim N(0,\tau^{-1})$, $\mathrm{E}\,|s|=2^{1/2}\pi^{-1/2}\cdot\tau^{-1/2}$ and $\mathrm{E}\,|s|^3=2^{3/2}\pi^{-1/2}\tau^{-3/2}$, we obtain the result.

3. Approximation to the covariance matrix

By the statement that a sequence of matrices, say B_n, approximates the covariance matrix of $\hat{\beta}$ up to $O(n^{-(r-1)/2})$, we shall mean that the covariance matrix $\mathrm{Cov}\,(\sqrt{n}\,\hat{\beta})=\mathrm{Cov}\,(d)=\mathrm{E}\,(dd')$ of $d=\sqrt{n}\,(\hat{\beta}-\beta)$ exists and B_n satisfies

$$\lim_{n\to\infty} n^{(r-1)/2}\|\mathrm{E}\,(dd')-B_n\|=0\ , \tag{3.1}$$

where for a matrix C, $\|C\|=(\mathrm{tr}\,CC')^{1/2}$. Hence, in this case the δ-method discussed in Section 2 is effective if the existence of $\mathrm{E}\,(dd')$ is verified and $B_n=\mathrm{E}\left(\sum_{i=0}^{r-1}d_i\right)\left(\sum_{i=0}^{r-1}d_i\right)'$ is shown to satisfy (3.1), where $\sum_{i=0}^{r-1}d_i$ is defined by (2.1). But in this approach, it is usually difficult to show (3.1). This point is discussed in Taylor [16]. On the other hand, in his problem Srivastava [13] called $\mathrm{E}\left(\sum_{i=0}^{r-1}d_i\right)\left(\sum_{i=0}^{r-1}d_i\right)'$ the *asymptotic covariance matrix in probability* up to $O(n^{-(r-1)/2})$ where $r=2$, which implicitly implied that neither the existence of $\mathrm{E}\,(dd')$ nor (3.1) had been checked.

In our approach, we use the relation

$$\mathrm{Cov}\,(d)=\mathrm{E}\,(H)=A^{-1}+\mathrm{E}\,(J)=A^{-1}+\mathrm{E}\,(J_1)+\mathrm{E}\,(R_1)\ . \tag{3.2}$$

Assuming $\mathrm{E}\,(H)$ exists, this implies

$$\|\mathrm{Cov}\,(d)-A^{-1}-\mathrm{E}\,(J_1)\|=\|\mathrm{E}\,(R_1)\|\ . \tag{3.3}$$

Thus if $\|\mathrm{E}\,(R_1)\|=O(n^{-3/2})$ is shown, $A^{-1}+\mathrm{E}\,(J_1)$ approximates $\mathrm{Cov}\,(d)=\mathrm{E}\,(H)$ at least up to $O(n^{-1})$. Further, evaluating $\|\mathrm{E}\,(R_1)\|$ gives a bound for the left side of (3.3). In the next section, we use

LEMMA 3.1. *Assume that* $\mathrm{E}\,(H)$ *exists and that all the* 3*rd*, 5*th and* 7*th moments of* $\bar{V}$ *in* (2.6) *are zero. Then* $\mathrm{Cov}\,(d)=A^{-1}+\mathrm{E}\,(J_1)+\mathrm{E}\,(J_2)+\mathrm{E}\,(R_2)$, *where*

(3.4) $$J_2=A^{-1/2}[D^2G+GD^2+DGD]A^{-1/2}\ , \qquad and$$
$$R_2=A^{-1/2}[GD^4Q+QD^3GD+DGD^3Q+QD^2GD^2Q]A^{-1/2}\ ,$$

and D, G and Q are defined in (2.6).

PROOF. The proof is given in Appendix A.

4. Approximations to the moments and pdf of the UZE

In this section, in the 2 equations SUR model we derive an approximate covariance matrix of the UZE with a bound of order n^{-3}, and an approximate pdf with a bound of order n^{-2}. The SUR model has the structure: in (1.1)

(4.1) $$X=\begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}, \quad y=\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad u=\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad \text{and} \quad \beta=\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

where $X_i: m\times k_i$, $y_i: m\times 1$, $u_i: m\times 1$, $\beta_i: k_i\times 1$, $k_1+k_2=k$ and $n=2m$. Here the covariance matrix of u is of the form

(4.2) $$\Omega=\mathrm{Cov}\,(u)=\Sigma\otimes I_m \qquad \text{where } \Sigma=(\sigma_{ij})\in S(2)\ .$$

By the UZE we mean the 2SAE, $\hat\beta(\hat\Omega)$ with $\hat\Omega=S\otimes I_m$ where with $Y=[y_1, y_2]$ and $X^*=[X_1, X_2]$, $S=Y'[I-X^*(X^{*\prime}X^*)^+X^{*\prime}]Y$. Let l be the rank of X^* and $q=m-l$. Then as is well known, with $\hat\Omega=S\otimes I$, $\hat\beta(\hat\Omega)$ satisfies the assumption (1.3) and $S\sim W_2(\Sigma, q)$, the Wishart distribution with mean $\mathrm{E}\,(S)=q\Sigma$ and d.f. q. By (2.3), transforming X and $S\otimes I$ with $\Omega^{-1/2}=\Sigma^{-1/2}\otimes I$ into

(4.3) $$\bar X=[\Sigma^{-1/2}\otimes I]XA^{-1/2}/\sqrt{n} \quad \text{and} \quad P=\Sigma^{-1/2}S\Sigma^{-1/2}\otimes I\equiv\tilde S\otimes I$$

the results in Sections 2 and 3 hold. Recall that $A=X'[\Sigma^{-1}\otimes I]X/n$ and $J=H-A^{-1}$ where H is given by (2.4).

We first show the existence of some moments of $d=\sqrt{n}\,[\hat\beta(S\otimes I)-\beta]$. To show the existence of moments of H, without loss of generality, we replace P^{-1} in H by

(4.4) $$P_0^{-1}=W\otimes I \qquad \text{where } W=2\tilde S^{-1}/\mathrm{tr}\,\tilde S^{-1}\ ,$$

so that $\mathrm{tr}\,W=2$. A convenience in the 2 equations SUR model is that W can be written as $W=2\tilde S^*/\mathrm{tr}\,\tilde S^*$, where

(4.5) $$\tilde S^*=\begin{bmatrix} \tilde s_{22} & -\tilde s_{11} \\ -\tilde s_{12} & \tilde s_{11} \end{bmatrix} \qquad \text{when } S=\begin{bmatrix} \tilde s_{11} & \tilde s_{12} \\ \tilde s_{21} & \tilde s_{22} \end{bmatrix}\ .$$

Since $\tilde S\equiv\Sigma^{-1/2}S\Sigma^{-1/2}\sim W(I_2, q)$, it is easy to see that $\tilde S^*\sim W(I_2, q)$. As in (2.6), letting $\bar V=W\otimes I-I$, $D=\bar X'\bar V\bar X$ and $E=\bar X'\bar V^2\bar X$ yields H of the form (2.7). Some moments of $W-I$, are listed in Appendix B. Further, let w_1 and w_2 be the latent roots of W where $w_1 \leqq w_2$. Then, $\operatorname{tr} W=2$ with $w_1 \leqq w_2$ implies $w_1+w_2=2$ and $1 \leqq w_2 \leqq 2$.

LEMMA 4.1. *The moments of* $d=\sqrt{n}[\hat{\beta}(S\otimes I)-\beta]$ *or* $\hat{\beta}(S\otimes I)$ *at least up to order* $q-3$ *exist.*

PROOF. Let $Z=w_2-1$. Then $0 \leqq Z \leqq 1$ and

$$H \leqq (1-Z^2)^{-1}A^{-1} \tag{4.6}$$

(Kariya [7]) and $Z^2 \sim Be(1, (1/2)(q-1))$, the beta distribution with d.f. 1 and $(1/2)(q-1)$. From (4.6), $t'Ht \leqq (1-Z^2)^{-1}t'A^{-1}t$ for all t. Hence $\mathrm{E}\,(t'Ht)^j \leqq (t'A^{-1}t)^j\,\mathrm{E}\,(1-Z^2)^{-j}$. Since $Z^2 \sim Be(1, (1/2)(q-1))$, $\mathrm{E}\,(1-Z^2)^{-j}$ exists if $j<(1/2)(q-1)$. Therefore, for $q>2j+1$, $\mathrm{E}\,(t'Ht)^j$ exists for all t. Hence the characteristic function $\phi(t)$ of d is expressed as $\phi(t)=\sum_{j=0}^{\alpha}\frac{1}{j!}\mathrm{E}\,(-(1/2)t'Ht)^j+O_\alpha(t)$, where α is the largest integer such that $\alpha<(1/2)(q-1)$ and $O_\alpha(t)$ satisfies $\lim_{t\to 0}|O_\alpha(t)/\|t\||=0$ (see, e.g. Breiman [2], p. 237). This implies that the moments of d exist up to $2\alpha<q-1$. Hence if q is even, the $(q-2)$th moments of d exist and if q is odd, the $(q-3)$th moments of d exist.

We are now in a position to apply the results in Sections 2 and 3. Let

$$\bar{X}=\begin{bmatrix}\bar{X}_1\\ \bar{X}_2\end{bmatrix}\begin{matrix}m\\ m\end{matrix}, \quad M=\bar{X}_1'\bar{X}_1-\bar{X}_2'\bar{X}_2 \quad \text{and} \quad N=\bar{X}_2'\bar{X}_1+\bar{X}_1'\bar{X}_2\,. \tag{4.7}$$

Recall that J_1, J_2 and R_2 are defined by (2.8) and (3.4) respectively.

THEOREM 4.1. *Let* $q>3$. *The covariance matrix of* $d=\sqrt{n}[\hat{\beta}(S\otimes I)-\beta]$ *is evaluated as* $\operatorname{Cov}(d)=A^{-1}+\mathrm{E}\,(J_1)+\mathrm{E}\,(J_2)+\mathrm{E}\,(R_2)$, *where*

$$\mathrm{E}\,(J_1)=\frac{2}{q-1}A^{-1}-\frac{1}{q-1}A^{-1/2}(M^2+N^2)A^{-1/2} \tag{4.8}$$

$$\begin{aligned}\mathrm{E}\,(J_2)=\frac{3}{(q+1)(q+3)}A^{-1/2}[&4(M^2+N^2)-2(M^4+N^4)\\ &-(M^2+N^2)^2-(MN+NM)^2]A^{-1/2}\,.\end{aligned} \tag{4.9}$$

Further, a bound for $\|\mathrm{E}\,(R_1)\|=\|\operatorname{Cov}(d)-A^{-1}-\mathrm{E}\,(J_1)\|$ *is given by*

$$\|\mathrm{E}\,(R_1)\| \leqq \frac{3.65}{(q+1)(q+3)}\|A^{-1}\| \qquad (q \geqq 8) \tag{4.10}$$

and a bound for $\|\mathrm{E}\,(R_2)\|=\|\operatorname{Cov}(d)-A^{-1}-\mathrm{E}\,(J_1)-\mathrm{E}\,(J_2)\|$ *is given by*

$$\|\mathrm{E}(R_2)\| \leqq \frac{79}{(q+5)(q+3)(q+1)}\|A^{-1}\| \qquad (q \geqq 8)\,. \tag{4.11}$$

PROOF. The proofs of (4.8) and (4.9) are given in Appendix C and the proofs of (4.10) and (4.11) are given in Appendix D.

We note that A^{-1}/n is the covariance matrix of the Gauss-Markov estimator $\bar{\beta} \equiv \hat{\beta}(\Sigma \otimes I) = (x'[\Sigma^{-1} \otimes I]X)^{-1}X'[\Sigma^{-1} \otimes I]y$.

Using the definition of $\bar{X}$, M and N in (4.3) and (4.7), it is easy to see that the approximate covariance matrix $A^{-1}+\mathrm{E}(J_1)$ with $\mathrm{E}(J_1)$ in (4.8) is equivalent to the asymptotic covariances in probability derived by Srivastava [13] where the RZE is treated. Theorem 4.1 not only gives the bound (4.10) for this approximation, but also it provides the higher order approximation $A^{-1}+\mathrm{E}(J_1)+\mathrm{E}(J_2)$ with the bound (4.11). The additional information $\mathrm{E}(J_2)$ on $\mathrm{Cov}(d)$ is necessary for a further comparison between the RZE and UZE, since the asymptotic covariance matrices up to $O(n^{-1})$ of these estimates are the same, as pointed out by Srivastava [13].

Next, based on Theorem 2.1 we shall derive an approximate pdf of d.

THEOREM 4.2. (i) *Let $f(x)$ be the pdf of $d=\sqrt{n}\,[\hat{\beta}(S \otimes I)-\beta]$ and let*

$$f_0(x) = \left[1+\frac{1}{q+1}(\alpha_1-\alpha_2(x))\right]\phi_A(x) \tag{4.12}$$

where

$$\alpha_1 = k - \mathrm{tr}\,(M^2+N^2)/2\,, \qquad \alpha_2(x) = x'Ax - x'A^{1/2}(M^2+N^2)A^{1/2}x \tag{4.13}$$

and

$$\phi_A(x) = (2\pi)^{-k/2}|A|^{1/2}\exp\left(-(1/2)x'Ax\right)\,. \tag{4.14}$$

Then

$$\sup_x |f(x)-f_0(x)| \leqq (2\pi)^{-k/2}2^{-3}|A|^{1/2}\left[\frac{4k(k+2)}{(q-3)^2}+\frac{15k}{(q+1)(q+3)}\right]\,. \tag{4.15}$$

(ii) *For normal approximation we obtain*

$$\sup_x |f(x)-\phi_A(x)| \leqq (2\pi)^{-k/2}2^{-1}\left[\frac{2k}{q-3}\right]|A|^{1/2}\,. \tag{4.16}$$

PROOF. Substituting $\mathrm{E}(J_1)$ in (4.8) into (2.9) yields (4.12). For (4.15), we use (2.10). From the proof of (4.10) in Appendix D, $\bar{R}_1 = \mathrm{E}(R_1) \leqq [3.57/(q+3)(q+1)]A^{-1}$, while from (4.13), $J=H-A^{-1} \leqq [2/(q-3)]\cdot$

A^{-1}. Therefore, the bound in (2.10) is further bounded above by

$$c(k, 2)\left[\frac{4}{(q-3)^2}\int (t'A^{-1}t)^2 \exp(-(1/2)t'A^{-1}t)dt + \frac{4\times 3.65}{(q+1)(q+3)}\int t'A^{-1}t \exp(-(1/2)t'A^{-1}t)dt\right].$$

Evaluating this yields (4.15). The proof of (4.16) is similar to that of (4.10).

The result (4.16) is comparable to a result in Zellner [19], in which when $q\geqq 20$ and $X_1'X_2=0$, normal approximation is suggested. To see this, for example, take $k=4$ and $q=23$, then the relative error is $\sup|f(x)-\phi_A(x)|/\phi_A(0)=1/5$. While the relative error based on (4.15) is $\sup|f(x)-f_0(x)|/f_0(0)<1/8[4k(k+2)/(q-3)^2+15k/(q+1)(q+3)]=0.042$. Hence, when $X_1X_2\neq 0$ and $q\leqq 23$ normal approximation may not be appropriate.

Finally, based on Theorem 2.2 (ii), we consider approximations to the cdf of a linear combination of d.

THEOREM 4.3. (i) *$G(Z)$ be the cdf of $z=a'd=\sqrt{n}\,a'[\hat{\beta}(S\otimes I)-\beta]$ and let $\tau=a'A^{-1}a$. Further, let $g_0(z)=[1+(1/2)(\tau^{-1}-z^2\tau^{-2})\,\mathrm{E}\,(a'J_1a)]\phi_\tau(z)$, $G_0(z)=\int_{-\infty}^{z} g_0(x)dx$, and $\phi_\tau(z)=(2\pi\tau)^{-1/2}\exp(-z^2/2\tau)$. Then*

$$\sup_z |G(z)-G_0(z)|\leqq 2\pi^{-1}\left[\frac{2}{(q-3)^2}+\frac{4}{(q+1)(q+3)}\right]. \tag{4.17}$$

(ii) *Let $\Phi_\tau(z)=\int_{-\infty}^{z}\phi_\tau(x)dx$. Then*

$$\sup_z |G(z)-\Phi_\tau(z)|\leqq 2^{1/2}\pi^{-1/2}(q-3)^{-1}. \tag{4.18}$$

PROOF. (i) First note that from Lemma 2.1 and (4.6), we obtain $\mathrm{E}\,(a'Ja)^2\leqq[2/(q-3)]^2\tau^2$ and from the proof of Theorem 4.2, $\mathrm{E}\,(a'R_1a)\leqq [3.75/(q+3)(q+1)]\tau$. Substituting these upper bounds into the bound in (2.14) yields the result. (ii) Immediate from Theorem 2.2 (ii).

We first remark that Philips [10] approximated $G(z)$ by $G_0(z)$ in a neighbourhood of $z=0$. But our result (4.17) holds uniformly for all z. Second, the approximation to the pdf of z can also be evaluated based on Theorem 2.2 (i), but it is omitted here.

Appendix A. Proofs of Lemmas 2.2 and 3.1.

Using the relations $DQ=QD$, $G=E-D^2$ and $Q=(I+D)^{-1}=I-DQ=I-D+D^2Q$, repeatedly, we can compute, $A^{1/2}HA^{1/2}$ as

$$(I+D)^{-1}(I+2D+E)(I+D)^{-1}$$
$$=(I-DQ)(I+2D+E)(I-DQ)$$
$$=I+E-D^2+2D^3-DE-ED+GD^2Q+QD^2G+QDGDQ\ .$$

Substituting this into $J=H-A^{-1}=A^{-1/2}(A^{1/2}HA^{1/2}-I)A^{-1/2}$, we obtain $J=J_1+R_1$ where J_1 and R_1 are defined by (2.8). This proves Lemma 2.2. Since $GD^2Q=GD^2-GD^3+GD^4Q$, $QD^2G=D^2G-D^3G+QD^4G$ and $QDGDQ=DGD-D^2GD-DGD^2+QD^3GD+DGD^3Q+QD^2GD^2Q$, substituting these into R_1 yields

$$R_1=A^{-1/2}(D^2G+GD^2+DGD)A^{-1/2}+A^{-1/2}(2D^3-DE-ED-GD^3-D^3G)A^{-1/2}$$
$$+A^{-1/2}(GD^4Q+QD^3GD+DGD^3Q+QD^2GD^2Q)A^{-1/2}=J_2+J_3+R_2\ .$$

But by assumption, $\mathrm{E}(J_3)=0$ and so $\mathrm{E}(R_1)=\mathrm{E}(J_2)+\mathrm{E}(R_2)$, completing the proof of Lemma 3.1.

Appendix B. Moments of $V\equiv W-I$.

For $W=2\tilde{S}^*/\mathrm{tr}\,\tilde{S}^*$ with $\tilde{S}^*$ in (4.4), let $W=(w_{ij})$ and $W-I\equiv V=(v_{ij})$. Then $v_{ii}=w_{ii}-1$ $(i=1, 2)$ and $v_{12}=w_{12}$. Since $w_{11}+w_{22}=2$, $v_{11}+v_{22}=0$ or $v_{11}=-v_{22}$. Since $\tilde{S}^*\sim W(I_2, q)$, we write $\tilde{S}^*$ in terms of normal variate $\tilde{Z}=[\tilde{Z}_1, \tilde{Z}_2]: q\times 2\sim N(0, I_q\otimes I_2)$ as

$$\tilde{S}^*=\tilde{Z}'\tilde{Z}=(\tilde{Z}_i'\tilde{Z}_j) \qquad (i, j=1, 2)\ . \tag{B.1}$$

Then from $W=2\tilde{S}^*/\mathrm{tr}\,\tilde{S}^*$, $w_{ij}=2\tilde{Z}_i'\tilde{Z}_j/(\tilde{Z}_1'\tilde{Z}_1+\tilde{Z}_2'\tilde{Z}_2)$ $(i, j=1, 2)$. Let $T_1=\tilde{Z}_1'\tilde{Z}_1$, $T_2=\tilde{Z}_2'\tilde{Z}_1(\tilde{Z}_1'\tilde{Z}_1)^{-1}\tilde{Z}_1'\tilde{Z}_2$ and $T_3=\tilde{Z}_2'(I-\tilde{Z}_1(\tilde{Z}_1'\tilde{Z}_1)^{-1}\tilde{Z}_1']\tilde{Z}_2$, and define

$$U_i=T_i/(T_1+T_2+T_3) \qquad (i=1, 2, 3)\ . \tag{B.2}$$

Since $T_1\sim X^2(q)$, $T_2\sim X^2(1)$ and $T_3\sim X^2(q-1)$ independently,

$$(U_1, U_2, U_3)\sim D_3(q/2, 1/2, (q-1)/2)\ , \tag{B.3}$$

where $D_3(\alpha, \beta, \tau)$ denotes the three dimensional Dirichlet distribution with parameters α, β and τ. Hence the moment of $U_1^{r_1}U_2^{r_2}U_3^{r_3}$ is easily computed (see, e.g. Johnson and Kotz [6], p. 231). Using $w_{11}=2U_1$, $w_{12}^2=4U_1U_2$ and $v_{11}=-v_{22}$, we can easily evaluate the joint moments of v_{ij}. Denoting the 2nd and 4th order moments of $v_{11}^{\alpha_1}v_{12}^{\alpha_2}v_{22}^{\alpha_3}$ by $a(\alpha_1, \alpha_2, \alpha_3)$, we have $a(1, 1, 0)=0$, $a(1, 0, 1)=-r$, $a(0, 1, 1)=0$, $a(2, 0, 0)=r$, $a(0, 2, 0)=r$, $a(0, 0, 2)=r$, $a(1, 1, 2)=0$, $a(1, 2, 1)=-p$, $a(2, 1, 1)=0$, $a(2, 2, 0)=p$, $a(2, 0, 2)=3p$, $a(0, 2, 2)=p$, $a(3, 1, 0)=0$, $a(3, 0, 1)=-3p$, $a(0, 3, 1)=0$, $a(1, 3, 0)=0$, $a(1, 0, 3)=-3p$, $a(0, 1, 3)=0$, $a(4, 0, 0)=3p$, $a(0, 4, 0)=3p$, $a(0, 0, 4)=3p$, where $p=1/(q+3)(q+1)$ and $r=1/(q+1)$.

The next lemma shows that the odd order moments of v_{ij}'s vanish.

LEMMA B.1. *If* $\alpha_1+\alpha_2+\alpha_3$ *is odd, then* $\mathrm{E}(v_{11}^{\alpha_1}v_{12}^{\alpha_2}v_{22}^{\alpha_3})=0$.

PROOF. From the definition of v_{ij} and from (B.1), v_{ii} is an even function of $\tilde{Z}_1$ and $\tilde{Z}_2$ $(i=1, 2)$, and v_{12} is an odd function of $\tilde{Z}_1$ (or $\tilde{Z}_2$). Note $\tilde{Z}_i$'s$\sim$iid $N(0, I_q)$. Hence if α_2 odds, $\mathrm{E}(v_{11}^{\alpha_1}v_{12}^{\alpha_2}v_{22}^{\alpha_3})=0$ follows immediately. If α_2 is even so that $\beta=\alpha_1+\alpha_3$ is odds, from $v_{11}=-v_{22}$, $\mathrm{E}(v_{11}^{\alpha_1}\cdot v_{12}^{\alpha_2}v_{22}^{\alpha_3})=(-1)^{\alpha_3}\mathrm{E}(v_{11}^{\beta}v_{12}^{\alpha_2})$. On the other hand, from the symmetry of v_{ij}'s, it is easy to see that the distribution of (v_{11}, v_{12}) is the same as that of (v_{22}, v_{21}). Hence $\mathrm{E}(v_{11}^{\beta}v_{12}^{\alpha_2})=\mathrm{E}(v_{22}^{\beta}v_{21}^{\alpha_2})=-\mathrm{E}(v_{11}^{\beta}v_{12}^{\alpha_2})$, implying $\mathrm{E}(v_{11}^{\beta}\cdot v_{12}^{\alpha_2})=0$. This proves Lemma B.1.

For Appendix C, we compute some more moments. Let

(B.4) $e_{11}=v_{11}^2+v_{12}^2$, $e_{12}=e_{21}=v_{11}v_{12}+v_{12}v_{22}$, and $e_{22}=v_{12}^2+v_{22}^2$.

Using $a(\alpha_1, \alpha_2, \alpha_3)$ defined above, after some calculation we obtain

(B.5) $\mathrm{E}(e_{11})=\mathrm{E}(e_{22})=2/(q+1)$; $\mathrm{E}(e_{11}^2)=\mathrm{E}(e_{22}^2)=\mathrm{E}(e_{11}e_{22})=8/(q+3)(q+1)$,

(B.6) $\mathrm{E}(e_{12}^2)=\mathrm{E}(e_{11}e_{12})=\mathrm{E}(e_{22}e_{12})=0$; $\mathrm{E}(e_{12}v_{ij}e_{kk})=0$ $\quad (i, j, k=1, 2)$,

(B.7) $\mathrm{E}(e_{ij}v_{12}e_{km})=0$ $\quad (i, j, k, m=1, 2)$;
$\mathrm{E}(v_{12}^2e_{ii})=4/(q+3)(q+1)$ $\quad (i, j=1, 2)$,

(B.8) $\mathrm{E}(e_{ii}v_{ij}^2)=4/(q+3)(q+1)$ $\quad (i, j=1, 2)$;
$\mathrm{E}(v_{11}v_{12}e_{ii})=-4/(q+3)(q+1)$ $\quad (i, j=1, 2)$.

Appendix C. Proofs of (4.8) and (4.9).

Introduce the notation $C_{ij}=\bar{X}_i'\bar{X}_j$ where $\bar{X}_i$'s are defined in (4.7) and $C_{11}+C_{22}=I_k$. Then

(C.1) $D=\bar{X}'\bar{V}\bar{X}=v_{11}C_{11}+v_{12}C_{12}+v_{21}C_{21}+v_{22}C_{22}$ and

(C.2) $E=\bar{X}'\bar{V}^2\bar{X}=e_{11}C_{11}+e_{12}C_{12}+e_{21}C_{21}+e_{22}C_{22}$,

where e_{ij}'s are given by (B.4). Using Appendix B and the definition of M and N in (4.7), it follows that

(C.3) $\mathrm{E}(D^2)=(M^2+N^2)/(q+1)$; $\mathrm{E}(E)=2I_k/(q+1)$

(C.4) $\mathrm{E}(D^2E)=\mathrm{E}(ED^2)=4(M^2+N^2)/(q+3)(q+1)$

(C.5) $\mathrm{E}(D^4)=\{3M^4+M^2N^2+(MN)^2+MN^2M+NM^2N+(NM)^2$
$+N^2M^2+3N^4\}/(q+3)(q+1)$

(C.6) $\mathrm{E}(DED)=4(M^2+N^2)/(q+3)(q+1)$.

Now using (C.3)-(C.6), $\mathrm{E}(J_1)$ and $\mathrm{E}(J_2)$ can be easily evaluated as

(C.7) $\mathrm{E}(J_1)=A^{-1/2}\,\mathrm{E}[E-D^2]A^{-1/2}=(4.8)$

(C.8) $\mathrm{E}(J_2)=A^{-1/2}\,\mathrm{E}[D^2G+GD^2+DGD]A^{-1/2}$
$=A^{-1/2}\,\mathrm{E}[D^2E+ED^2+DED-3D^4]A^{-1/2}=(4.9)$.

Further, since J_3 in Appendix A consists of only odd (joint) moments of v_{ij}'s, $\mathrm{E}(J_3)=0$ follows from Appendix B.

Appendix D. Proofs of (4.10) and (4.11).

Let ρ be a 2×2 orthogonal matrix such that $\rho W\rho'=\mathrm{diag}\{w_1, w_2\}=$ diagonal matrix with diagonal elements w_1 and w_2 in this order and let

(D.1) $$[\rho\otimes I]\bar{X}=\tilde{X}=\begin{bmatrix}\tilde{X}_1\\ \tilde{X}_2\end{bmatrix}\begin{matrix}m\\ m\end{matrix} \quad \text{and} \quad B=\tilde{X}_2'\tilde{X}_2 .$$

Then $\mathrm{tr}\,W=2$ with $w_1\leqq w_2$ implies $w_1+w_2=2$ and $1\leqq w_2\leqq 2$ and $\bar{X}'\bar{X}=I_k$ implies $I_k=\tilde{X}'\tilde{X}=\tilde{X}_1'\tilde{X}_1+\tilde{X}_2'\tilde{X}_2$. It is noted that ρ and so $\tilde{X}$ are random. Further, let $0\leqq\delta_1\leqq\cdots\leqq\delta_k$ be the latent roots of $B=\tilde{X}_2'\tilde{X}_2$ and let Ψ be a $k\times k$ orthogonal matrix such that

(D.2) $$\Psi B\Psi'=\mathrm{diag}\{\delta_1,\cdots,\delta_k\} ,$$

Since $B=I-\tilde{X}_1'\tilde{X}_1$, $0\leqq\delta_i\leqq 1$ $(i=1,\cdots,k)$. Note $Z=w_2-1$ and $0\leqq Z<1$. Using $\tilde{X}_1'\tilde{X}_1=I-B$ and $w_1=2-w_2$, D, E, G and Q in (2.14) with $\bar{V}=W\otimes I-I$ as

(D.3) $$D=Z(2B-I)\,,\quad E=Z^2I\,,\quad G=4Z^2B(I-B)\,,$$
$$Q=(I+Z(2B-I))^{-1} .$$

It is easy to see that D, E, G and Q commute mutually. Now we prove (4.10). Since $\mathrm{E}(D^3)=0$, $\mathrm{E}(DE)=0$ and $\mathrm{E}(ED)=0$ (see Appendix B), from (2.8) $A^{1/2}\,\mathrm{E}(R_1)A^{1/2}$ is equal to

(D.4) $$\mathrm{E}\{GD^2Q+QD^2G+QDGDQ\}=\mathrm{E}\{3GD^2+4GD^2Q+D^4GQ^2\} ,$$

where $QD^2G=GD^2Q=GD^2-GD^3+GD^4Q$ and $QDGDQ=D^2G-2D^3G+2D^4GQ+D^4GQ^2$ with $\mathrm{E}(GD^3)=\mathrm{E}(D^3G)=0$ are used. From (D.3) $GD^2=4Z^4B(I-B)(2B-I)$, and hence from (D.2), $\Psi GD^2\Psi'=4Z^2\,\mathrm{diag}\{h(\delta_1),\cdots,h(\delta_k)\}$ where $h(\delta)=\delta(1-\delta)(2\delta-1)^2$ $(0\leqq\delta\leqq 1)$. But $h(\delta)$ is maximized at $\delta_0=(2\pm\sqrt{2})/4$ and $h(\delta_i)\leqq h(\delta_0)=1/4^3$ $(i=1,\cdots,k)$. Therefore $GD^2\leqq 4Z^2h(\delta_0)I$. Since from Lemma 4.1, $Z^2\sim Be(1,(1/2)(q-1))$,

(D.5) $$\mathrm{E}(GD^2)\leqq[(1/2)(q+3)(q+1)]I\equiv\gamma_0 I .$$

Similarly, $GD^4Q=4Z^6B(I-B)(2B-I)^4(I+Z(2B-I))^{-1}=4Z^6\Psi'\,\mathrm{diag}\{h(\delta_1),\cdots,h(\delta_k)\}\Psi$ where in this case $h(\delta)=\delta(1-\delta)(2\delta-1)^6(1+Z(2\delta-1))^{-1}$. Since

$4Z^6h(\delta)\leqq 2Z^5(1-\delta)(2\delta-1)^4$ from $2Z\delta\leqq 1-Z+2Z\delta$, with $\delta_0=9/10$, $4Z^6h(\delta)\leqq 2Z^5(1-\delta_0)(2\delta_0-1)^4\leqq 2Z^4(1-\delta_0)(2\delta_0-1)^4$. Therefore

(D.6) $\quad \mathrm{E}\,(GD^4Q)\leqq\gamma_1 I\leqq(4/5)^5[1/(q+3)(q+1)]I\,,\qquad$ where

(D.7) $\quad \gamma_1=(1/5)(4/5)^4 15\sqrt{\pi}\,\Gamma((q+1)/2)/2(q+4)(q+2)\Gamma((q+2)/2)\,.$

Next, we evaluate $\mathrm{E}\,(D^4GQ^2)$. Note $D^4GQ^2=4Z^6\Psi'\operatorname{diag}\{h(\delta_1),\cdots,h(\delta_k)\}\Psi$ where $h(\delta)=\delta(1-\delta)(2\delta-1)^4(1+Z(2\delta-1))^{-2}$. Hence from $4Z^6h(\delta)\leqq Z^5(1-Z^2)^{-1}(1-\delta)(2\delta-1)^4\leqq Z^5(1-Z^2)^{-1}(1/10)(4/5)^4\leqq Z^4(1-Z^2)^{-1}(1/10)(4/5)^4$ where $4Z\delta(1+Z(2\delta-1))^{-2}\leqq(1/2)(1-Z)^{-1}\leqq 1/(1-Z^2)$ is used.

(D.8) $\quad \mathrm{E}\,(D^4GQ^2)\leqq\gamma_2 I\leqq(1/10)(4/5)^4 8[1/(q+1)(q-3)]/I\,,\qquad$ where

(D.9) $\quad \gamma_2=(1/10)(4/5)^4 15\sqrt{\pi}\,\Gamma((q+1)/2)/2(q+2)(q-3)\Gamma((q+2)/2)\,.$

Thus, from (D.4), (D.5), (D.6) and (D.8)

(D.10) $\quad \|\mathrm{E}\,(R_1)\|\leqq(3\gamma_0+4\gamma_1+\gamma_2)\|A^{-1}\|\leqq\dfrac{2.82}{(q+3)(q+1)}+\dfrac{0.33}{(q+1)(q-3)}\,.$

If $q\geqq 8$, the right side of (D.10) is bounded above by $3.65/(q+3)(q+1)$. This proves (4.10).

Secondly, we prove (4.11). From (3.5) and (D.6), it is necessary to evaluate $\mathrm{E}\,(QD^2GD^2Q)$, and $\mathrm{E}\,(D^4GQ)$. Note from $Q=I-D+D^2Q$

(D.11) $\quad \mathrm{E}\,(QD^2GD^2Q)=\mathrm{E}\,[D^4G+D^6G+2D^6GQ-2D^7GQ+D^8GQ^2]\,.$

In a similar manner, each term is evaluated as

(D.12) $\quad \mathrm{E}\,(D^4G)\leqq(2^6/3^2)[(q+5)(q+3)(q+1)]^{-1}I=\gamma_3 I\,,$

(D.13) $\quad \mathrm{E}\,(D^6G)\leqq(3^4/2)[(q+7)(q+5)(q+3)(q+1)]^{-1}I\equiv\gamma_4 I\,,$

(D.14) $\quad -\mathrm{E}\,(D^7GQ)=\mathrm{E}\,(D^8GQ)\leqq 3.2^4(8/9)^9[(q+7)(q+5)(q+3)(q+1)]^{-1}I\equiv\gamma_5 I\,,$

(D.15) $\quad \mathrm{E}\,(D^6GQ)=\mathrm{E}\,(D^6G)+\mathrm{E}\,(D^8GQ)\leqq(\gamma_4+\gamma_5)I\equiv\gamma_6 I\,,$

(D.16) $\quad \mathrm{E}\,(D^8GQ)\leqq 3.2^3(8/9)^9[(q+5)(q+3)(q+1)(q-3)]^{-1}I\equiv\gamma_7 I\,.$

On the other hand, $\mathrm{E}\,(D^4GQ)=\mathrm{E}\,(D^4G)+\mathrm{E}\,(D^6GQ)$, hence

(D.17) $\quad \mathrm{E}\,(D^4GQ)\leqq(\gamma_3+\gamma_6)I\,.$

Therefore $\|\mathrm{E}\,(R_2)\|\leqq[3(\gamma_3+\gamma_6)+\gamma_3+\gamma_4+2\gamma_5+2\gamma_6\gamma_7]\|A^{-1}\|$, which is bounded above by 79η where $\eta=[(q+5)(q+3)(q+1)]^{-1}$. Here it is used that $9\gamma_3\leqq 64\eta$, $\gamma_4\leqq 2.9\eta$, $7\gamma_5\leqq 10\eta$ and $\gamma_7\leqq 2.1\eta$.

HITOTSUBASHI UNIVERSITY
HIROSHIMA UNIVERSITY

REFERENCES

[1] Bhattacharya, R. N. and Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion, *Ann. Statist.*, **6**, 434-451.

[2] Breiman, L. (1968). *Probability*, Addison-Wesley, New York.

[3] Durbin, J. (1980). Approximations for densities of sufficient estimates, *Biometrika*, **67**, 311-333.

[4] Feller, W. (1966). *An Introduction to Probability and its Applications*, Vol. 2, Wiley, New York.

[5] Hall, A. (1976). The relative efficiency of estimates of seemingly unrelated regressions, Ph.D. Thesis, London School of Economics.

[6] Johnson, N. and Kotz, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions*, Wiley, New York.

[7] Kariya, T. (1981). Efficiencies of Zellner's estimator in the SUR model and a 2SAE in a heteroscedastic model, *J. Amer. Statist. Ass.*, **76**, 975-979.

[8] Kunitomo, N. (1977). A note on the efficiency of Zellner's estimator for the case of two seemingly unrelated regression equations, *Econ. Studies Quarterly*, **28**, 73-77.

[9] Philips, P. C. B. (1977). A general theorem in the theory of asymptotic expansions as approximations to the finite sample distributions of econometric estimators, *Econometrica*, **45**, 1517-1534.

[10] Philips, P. C. B. (1977a). An approximation to the finite sample distribution of Zellner's seemingly unrelated regression estimator, *J. Econometrics*, **6**, 147-164.

[11] Revanker, N. S. (1976). Use of restricted residuals in SUR systems: some finite sample results, *J. Amer. Statist. Ass.*, **71**, 183-188.

[12] Sargan, J. D. (1975). Gram-Charlier approximations applied to t-ratios of k-class estimators, *Econometrica*, **43**, 327-346.

[13] Srivastava, V. K. (1970). The efficiency of estimating seemingly unrelated regression equations, *Ann. Inst. Statist. Math.*, **22**, 483-493.

[14] Srivastava, V. K. and Dwivedi, T. D. (1979). Estimation of seemingly unrelated regression equations: A brief survey, *J. Econometrics*, **10**, 15-32.

[15] Srivastava, V. K. and Upadhyaya, S. (1978). Large-sample approximations in seemingly unrelated regression equations, *Ann. Inst. Statist. Math.*, **30**, A, 89-96.

[16] Taylor, W. E. (1977). Small sample properties of a class of two stage Aitken estimators, *Econometrica*, **45**, 497-508.

[17] Taylor, W. E. (1978). The heteroscedastic linear model: Exact finite sample results, *Econometrica*, **46**, 663-675.

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

[19] Zellner, A. (1963). Estimators for seemingly unrelated regressions: some finite sample results, *J. Amer. Statist. Ass.*, **58**, 977-992.