Ann. Inst. Statist. Math.
39 (1987), Part A, 575-591

A THIRD ORDER OPTIMUM PROPERTY OF THE ML ESTIMATOR IN A LINEAR FUNCTIONAL RELATIONSHIP MODEL AND SIMULTANEOUS EQUATION SYSTEM IN ECONOMETRICS*

NAOTO KUNITOMO

(Received Apr. 30, 1985; revised July 30, 1986)

Summary

The maximum likelihood (ML) estimator and its modification in the linear functional relationship model with incidental parameters are shown to be third-order asymptotically efficient among a class of almost median-unbiased and almost mean-unbiased estimators, respectively, in the large sample sense. This means that the limited information maximum likelihood (LIML) estimator in the simultaneous equation system is third-order asymptotically efficient when the number of excluded exogenous variables in a particular structural equation is growing along with the sample size. It implies that the LIML estimator has an optimum property when the system of structural equations is large.

1. Introduction

The concept of asymptotic higher order efficiency of estimation has been recently developed by several statisticians. (Ghosh et al. [10], Pfanzagle and Wefelmeyer [17], and Akahira and Takeuchi [1], for instance.) According to this theory, the maximum likelihood (ML) estimator and the Bayesian estimator with proper priors have third-order asymptotic efficiency under some regularity conditions. This means that given an estimator we can always construct a modified ML estimator which has the same asymptotic bias and smaller asymptotic loss than the estimator to be compared in the regular case. The purpose of the

* The research was partly supported by National Science Foundation Grant SES 79-13976 at the Institute for Mathematical Studies in the Social Sciences, Stanford University and Grant-in-Aid 60301081 of the Ministry of Education, Science and Culture at the Faculty of Economics, University of Tokyo. This paper was originally written as a part of the author's Ph.D. dissertation submitted to Stanford University in August, 1981. Some details of the paper were deleted at the suggestion of the associate editor of this journal.
Key words and phrases: Maximum likelihood estimator, third-order efficiency, linear functional relationship, incidental parameter, LIML estimator, large econometric models.

present paper is to show that the ML estimator itself has a third-order optimum property among median-unbiased estimators in a linear functional relationship model or the errors-in-variable model, which is a well-known irregular case. We also show that a modification of the ML estimator has a third-order optimum property among almost mean-unbiased estimators.

Because the number of parameters increases together with the sample size (denoted by N) in linear functional relationship models, we cannot necessarily apply general theorems in the regular asymptotic theory to these statistical models. In fact, the least squares (LS) estimator is inconsistent while the ML estimator is consistent but does not attain the Crámer-Rao lower bound in this case. However, the ML estimator attains the lower bound of asymptotic variance, which is larger than the Crámer-Rao lower bound, among a certain class of consistent estimators (see Theorem 3.1 in Section 3). Therefore, further comparison of estimators should be made in terms of higher order terms in the asymptotic expansion of their distribution functions.

Anderson [3] first shed some light on connections between the estimation problem of linear functional relationships and that of structural equation in the simultaneous equation system in econometrics. The ML estimator of the slope in the linear functional relationship is mathematically equivalent to the limited information maximum likelihood (LIML) estimator of a structural coefficient when the covariance matrix of the reduced form is known in the simultaneous equation system, and the LS estimator in the former is equivalent to the two-stage least squares (TSLS) estimator in the latter. Further, Anderson [4] has shown that the parameter sequence in which the noncentrality parameter (or the spread of incidental parameters) increases while the sample size N stays fixed in the linear functional relationship model corresponds to the "large sample" asymptotic theory developed in econometrics. In the large sample asymptotic theory in econometrics, the LIML and TSLS estimators are best asymptotically normal (BAN) estimators, namely, these estimators are consistent and the estimators normalized by the square root of the sample size T (not N) have the same limiting joint normal distribution with the covariance of inverse of the standardized Fisher information matrix. Here we should distinguish the sample size T in the simultaneous equation system from the sample size N in the linear functional relationship model. Since two BAN estimators are available, several modifications of the LIML and TSLS estimators have been proposed in hoping that they may improve BAN estimators in some sense. Anderson, Kunitomo and Morimune [6] made a systematic comparison of these efforts in the regular asymptotic theory in econometrics. Moreover, Takeuchi and Morimune [19] has shown that the LIML estimator has a third-order optimum property in this framework where N is fixed and the problem of incidental parameters does not arise.

On the other hand, Kunitomo [11] argued that the asymptotic theory in which both the noncentrality (or the spread of incidental parameters) and the sample size N (not T) increase is more appropriate in linear functional relationship models. Anderson [4] has shown that the sample size N minus one is the number of excluded exogenous variables in the structural equation of interest in the simultaneous equation system, say, K_2. (Although Anderson [3] uses two endogenous variables case, his arguments hold in the general case. See Kunitomo [13], for instance.) Recent macro-econometric models are more or less large in their size and hence K_2 is fairly large even if there are only two endogenous variables in a particular structural equation. The above parameter sequence can be interpreted as a new asymptotic theory, called the large-K_2 asymptotics, for large econometric models (Kunitomo [12]).

As K_2 (or $N-1$) increases along with the sample size T, the LIML estimator is consistent and asymptotically efficient while both the TSLS and the ordinary least squares (OLS) estimators are inconsistent under appropriate conditions. Furthermore, the modifications of the LIML estimator by Fuller [9] and Morimune and Kunitomo [16] are shown to improve the LIML estimator in terms of the asymptotic mean squared error, which is defined by the mean squared error of the asymptotic expansion of their distributions. Hence there has been still some confusion on the higher order asymptotic optimality of estimator in this situation. Hopefully this paper will clarify this ambiguity. The results obtained in Section 2 imply that the LIML estimator is third-order efficient among a class of almost median-unbiased estimators and a modification of the LIML estimator is third-order asymptotically efficient among a class of almost mean-unbiased estimators. Thus, the LIML estimation method gives the best estimator if we adjust the asymptotic bias according to our choice of criterion: median-unbiasedness or mean-unbiasedness, etc.

One important approach studied in econometrics in the past is small sample theory. Anderson and Sawa [5] and Anderson, Kunitomo and Sawa [6] evaluated the exact distribution functions of the TSLS and LIML estimators, respectively, with a systematic choice of the nuisance key parameters in the simultaneous equation system. The most important finding in their studies is that the TSLS estimator is badly biased while the distribution of the LIML estimator is centered at the true parameter value when K_2 (or $N-1$) is large. In this respect, the results reported in Sections 2 and 3 give some theoretical support to their findings.

We shall present the model, the assumptions and the statement of main theorems in Section 2. A general model and some implications of our results in econometrics will be discussed in Section 3. Proofs of theorems are given in Section 4.

2. Main results

Suppose that (x_{gh}, y_{gh}) is an observation from a bivariate normal distribution with mean (μ_h, ν_h), the covariance matrix Σ for $g=1,\cdots,M$ $(M\geqq 1)$, $h=1,\cdots,N$ $(N>1)$ and that the observations are independent. In this notation M is the number of replications and N is the sample size. Incidental parameters (μ_h, ν_h) are assumed to satisfy a linear functional relationship:

$$\nu_h=\alpha+\beta\mu_h\,, \qquad h=1,\cdots,N\,. \tag{2.1}$$

The angle θ between the line (2.1) and the μ_h-axis may replace the slope coefficient by $\beta=\tan\theta$. It will be convenient to write

$$\begin{pmatrix} x_{gh} \\ y_{gh} \end{pmatrix}=\begin{pmatrix} \mu_h \\ \nu_h \end{pmatrix}+\begin{pmatrix} u_{gh} \\ v_{gh} \end{pmatrix}\,, \tag{2.2}$$

where u_{gh} and v_{gh} are normally distributed random variables with means zero and covariance matrix $\Sigma=\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$. In this section we first deal with the linear functional relationship (2.1) without replication $(M=1)$ and the covariance matrix $\Sigma=\sigma^2 I_2$, where σ^2 is the unknown variance parameter. Alternatively, we may assume $\Sigma=\sigma^2\Omega$, where Ω is known. This model can be reduced to the above case by making a transformation of (2.2). If Σ is unknown completely, it is well-known that the unknown parameters of the model are not identified except the case of replicated observations $(M>1)$. See Anderson [4] for this issue. In the next section we shall deal with more general models $(M>1)$ and discuss their relations with the simultaneous equation system in econometrics.

We are particularly interested in estimating the slope coefficient or its angle θ with μ-axis in (2.1). The estimator of α in any method in this paper is $\hat{\alpha}=\bar{y}-\hat{\beta}\bar{x}$, where $\hat{\beta}$ is an estimator of β, $\bar{x}=(1/N)\sum_{h=1}^{N} x_h$ and $\bar{y}=(1/N)\sum_{h=1}^{N} y_h$. We denote that $x_h=x_{gh}$ and $y_h=y_{gh}$ in this case. Any estimator of the angle parameter θ is defined by $\hat{\beta}=\tan\hat{\theta}$ $(|\theta|\leqq\pi/2)$ where $\hat{\beta}$ is an estimator of the slope coefficient β. Other notations used here are the standardized estimator $\hat{e}=\lambda(\hat{\beta}-\beta)/(1+\beta^2)$ for any estimator of β and the noncentrality parameter

$$\lambda^2 = \frac{(1+\beta^2)}{\sigma^2} \sum_{h=1}^{N} (\mu_h - \bar{\mu})^2 , \tag{2.3}$$

where $\bar{\mu} = (1/N) \sum_{h=1}^{N} \mu_h$. The parameter λ^2 may be interpreted as a measure of the spread of incidental parameters $(\mu_1, \cdots, \mu_N)$ about their means. The assumption we shall make here to derive the asymptotic expansions of distributions of estimators is the following.

ASSUMPTION A. There exists a finite positive number ρ such that $\sum_{h=1}^{N} (\mu_h - \bar{\mu})^2/(n\sigma^2) = \rho + O(n^{-1})$ and $\delta = (1+\beta^2)\rho$, where $n = N-1$.

Assumption A means that the noncentrality parameter (or the spread of incidental parameters) increases with the same rate as the sample size N. To avoid complexity of expressions, λ^2 will be used instead of N since λ^2 is replaceable with $n\delta$ under Assumption A in the following analysis. It is possible to extend our results in this section to alternative parameter sequences instead of Assumption A, which will be discussed in Section 3.

Define a class of estimators of β by

$$\hat{\beta} = \phi(s_{yy}, s_{yx}, s_{xx}) + \frac{1}{n} \psi(s_{yy}, s_{yx}, s_{xx}) , \tag{2.4}$$

where

$$S = \begin{pmatrix} s_{xx} & s_{xy} \\ s_{xy} & s_{yy} \end{pmatrix} = \frac{1}{n} \sum_{h=1}^{N} \begin{pmatrix} x_h - \bar{x} \\ y_h - \bar{y} \end{pmatrix} \begin{pmatrix} x_h - \bar{x} \\ y_h - \bar{y} \end{pmatrix}' ,$$

where ϕ is four times continuously differentiable, ψ is twice continuously differentiable, and both functions are independent of N. We assume that all of these derivatives evaluated at true parameter values are bounded. This class is sometimes called the extended regular estimators (see Akahira and Takeuchi [1]). It includes the ML, the least squares (LS), and the generalized least squares (GLS) estimators, and their modifications.

Let also define the k-th order asymptotic median-unbiased (AMDU) estimator by

$$\limsup_{n\to\infty} \sup_{\beta \in U_\varepsilon} n^{(k-1)/2} \left| \Pr(\hat{\beta} \leqq \beta) - \frac{1}{2} \right| = \limsup_{n\to\infty} \sup_{\beta \in U_\varepsilon} n^{(k-1)/2} \left| \Pr(\hat{\beta} \geqq \beta) - \frac{1}{2} \right| = 0 ,$$

where U_ε is a neighborhood $|\beta - \beta_0| \leqq \varepsilon$ for some $\varepsilon > 0$ and any β_0.

THEOREM 2.1. *For all* $\xi_1 \geqq 0$ and $\xi_2 \geqq 0$,

$$\lim_{n\to\infty} n(\Pr\{-\xi_1<\sqrt{n}(\hat{\beta}_{ML}-\beta)\leqq\xi_2\}-\Pr\{-\xi_1<\sqrt{n}(\hat{\beta}-\beta)\leqq\xi_2\})\geqq 0\ ,$$

where $\hat{\beta}$ is any third-order AMDU estimator and $\hat{\beta}_{ML}$ is given by

$$\hat{\beta}_{ML}=\frac{s_{yy}-s_{xx}+((s_{yy}-s_{xx})^2+4s_{xy}^2)^{1/2}}{2s_{xy}}\ . \tag{2.5}$$

COROLLARY 2.1. *The ML estimator has a third-order optimum property among AMDU estimators with respect to any bounded bowl-shaped loss function $L_n(\beta,\hat{\beta})=h(n^{1/2}(\hat{\beta}-\beta))$ whose minimum value is zero at $\hat{\beta}=\beta$ and which increases with $|\hat{\beta}-\beta|$.*

We also consider the class of k-th order asymptotic mean-unbiased (AMNU) estimator by

$$\lim_{n\to\infty}\sup_{\beta\in U_\epsilon} n^{(k-1)/2}|\mathrm{AM}_n(\hat{\beta}-\beta)|=0\ ,$$

where $\mathrm{AM}_n(\cdot)$ stands for the expectation operator with respect to the Edgeworth expansion of the distribution function of $\hat{\beta}$.

THEOREM 2.2. *For all $\xi_1\geqq 0$ and $\xi_2\geqq 0$,*

$$\lim_{n\to\infty} n(\Pr\{-\xi_1<\sqrt{n}(\hat{\beta}^*-\beta)\leqq\xi_2\}-\Pr\{-\xi_1<\sqrt{n}(\hat{\beta}-\beta)\leqq\xi_2\})\geqq 0\ ,$$

where $\hat{\beta}$ is any third-order AMNU estimator. $\hat{\beta}^*$ is given by

$$\hat{\beta}^*=\frac{2s_{xy}}{s_{xx}-s_{yy}+2\hat{c}+[(s_{xx}-s_{yy})^2+4s_{xy}^2]^{1/2}} \tag{2.6}$$

where $n\hat{c}=1+l_1/(l_2-l_1)$ and l_i $(i=1,2)$ are the smaller and larger characteristic roots of matrix S.

COROLLARY 2.2. *The estimator $\hat{\beta}^*$ has a third-order optimum property among AMNU estimators with respect to any bounded bowl-shaped loss function $L_n(\sqrt{n}(\hat{\beta}-\beta))$.*

Turning to the estimation of the angle parameter θ, we obtain the following result.

THEOREM 2.3. *For all $\xi_1\geqq 0$ and $\xi_2\geqq 0$,*

$$\lim_{n\to\infty} n(\Pr\{-\xi_1<\sqrt{n}(\hat{\theta}_{ML}-\theta)\leqq\xi_2\}-\Pr\{-\xi_1<\sqrt{n}(\hat{\theta}-\theta)\leqq\xi_2\})\geqq 0\ ,$$

where $\hat{\theta}$ is any third-order AMDU or AMNU estimator.

COROLLARY 2.3. *The ML estimator of the angle has a third-order optimum property among third-order AMDU and AMNU estimators with respect to any bounded bowl-shaped loss function $L_n(\sqrt{n}(\hat{\theta}-\theta))$.*

3. Asymptotic theories in linear functional relationship models and simultaneous equation system in econometrics

We now consider the linear functional relationship model (2.2) with replications ($M>1$) and an arbitrary covariance matrix $\boldsymbol{\Sigma}$. In this case the log-likelihood function $\log L(\cdot)$ times -2 is proportional to

$$
\begin{aligned}
(3.1)\qquad & MN\log[\sigma_1^2\sigma_2^2(1-\tau^2)]+\frac{1}{(1-\tau^2)}\sum_{g=1}^{M}\sum_{h=1}^{N}\left\{\left(\frac{x_{gh}-\mu_h}{\sigma_1}\right)^2\right.\\
&\left.-2\tau\left(\frac{x_{gh}-\mu_h}{\sigma_1}\right)\left(\frac{y_{gh}-\nu_h}{\sigma_2}\right)+\left(\frac{y_{gh}-\nu_h}{\sigma_2}\right)^2\right\},
\end{aligned}
$$

where we write $\boldsymbol{\Sigma}=\begin{pmatrix}\sigma_1^2 & \tau\sigma_1\sigma_2\\ \tau\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$. From the log-likelihood function (3.1), the information matrix for $(\beta, \alpha, \mu_1,\cdots,\mu_N)$ is

$$
(3.2)\qquad \frac{M}{(1-\tau^2)\sigma_2^2}\begin{pmatrix} \sum_{g=1}^{N}\begin{pmatrix}\mu_g\\ 1\end{pmatrix}(\mu_h\ \ 1) & \begin{pmatrix}\mu_1,\cdots,\mu_N\\ 1,\cdots,1\end{pmatrix}\left(\beta-\tau\frac{\sigma_2}{\sigma_1}\right)\\ \begin{pmatrix}\mu_1 & 1\\ \vdots & \vdots\\ \mu_N & 1\end{pmatrix}\left(\beta-\tau\frac{\sigma_2}{\sigma_1}\right) & I_N\left(\beta^2+\frac{\sigma_2^2}{\sigma_1^2}-2\tau\frac{\sigma_2}{\sigma_1}\right)\end{pmatrix}.
$$

Then the partial information for β is given by

$$
(3.3)\qquad I(\beta)=\frac{M}{(\sigma_{11}\beta^2-2\beta\sigma_{12}+\sigma_{22})}\sum_{g=1}^{N}(\mu_g-\bar{\mu})^2 .
$$

The noncentrality parameter λ^2 may be defined by $(\sigma_{11}\beta^2-2\beta\sigma_{12}+\sigma_{22})^2I(\beta)/|\boldsymbol{\Sigma}|$ in the general case. Here we note that the partial information for β when $\boldsymbol{\Sigma}$ is unknown is the same as (3.3) because $\partial^2\log L/\partial\beta\partial\sigma_{ij}=\partial^2\log L/\partial\alpha\partial\sigma_{ij}=\partial^2\log L/\partial\mu_k\partial\sigma_{ij}=0$ for $i, j=1, 2$ and $k=1,\cdots, N$.

Let $(\bar{x}_h, \bar{y}_h)'=(1/M)\sum_{g=1}^{M}(x_{gh}, y_{gh})'$ and $(\bar{x}, \bar{y})'=(1/N)\sum_{h=1}^{N}(\bar{x}_h, \bar{y}_h)'$. Then the estimator of β is based on two sum of squares:

$$
(3.4)\quad \boldsymbol{H}=M\sum_{h=1}^{N}\begin{pmatrix}\bar{x}_h-\bar{x}\\ \bar{y}_h-\bar{y}\end{pmatrix}\begin{pmatrix}\bar{x}_h-\bar{x}\\ \bar{y}_h-\bar{y}\end{pmatrix}', \qquad \boldsymbol{G}=\sum_{g=1}^{M}\sum_{h=1}^{N}\begin{pmatrix}x_{gh}-\bar{x}_h\\ y_{gh}-\bar{y}_h\end{pmatrix}\begin{pmatrix}x_{gh}-\bar{x}_h\\ y_{gh}-\bar{y}_h\end{pmatrix}',
$$

where the degrees of freedom of $\boldsymbol{H}$ and $\boldsymbol{G}$ are $N-1=n$ and $(M-1)N$, respectively. When $M=1$, $\boldsymbol{H}$ is reduced to $n\boldsymbol{S}$ in Section 2. In this case we assumed that $\boldsymbol{\Sigma}$ is proportional to the identity matrix for identification of the model. When $M>1$, we have some extra degrees of freedom to estimate the covariance matrix $\boldsymbol{\Sigma}$ by $\boldsymbol{G}/(N(M-1))$. The ML estimator of β for $M>1$ corresponds to the characteristic vector $(\beta, -1)'$ associated with the smaller root of

$$\left|\frac{1}{n}\boldsymbol{H}-l\frac{1}{(M-1)N}\boldsymbol{G}\right|=0 . \tag{3.5}$$

Let a structural equation in time period t in a system of simultaneous equations be

$$y_{2t}=\beta y_{1t}+\sum_{k=1}^{K_1}\gamma_{1k}z_{kt}+u_{1t} , \qquad t=1,\cdots,T , \tag{3.6}$$

where two endogenous variables (y_{1t}, y_{2t}) and K_1 exogenous variables z_{1k} are appeared in this structural equation. In the simultaneous equation system there are many (at least two) endogenous variables and exogenous variables appeared in other structural equations. Anderson [3] and [4] clarified the relation between the linear functional relationship model with replications and the simultaneous equation system. As he pointed out that mathematically there is a one-to-one correspondence between two models with different terminologies. In the present notations, n $(=N-1)$ corresponds to K_2, which is the number of excluded exogenous variables in the particular structural equation (3.6). The degrees of freedom $(M-1)(n+1)$ correspond to $q=T-K$, where T is the sample size and K is the total number of exogenous variables in the simultaneous equation system. We note that $K=K_1+K_2$ and K_1 is the number of included exogenous variables in the particular structural equation (3.6). The partial information $I(\beta)$ in (3.3) corresponds to the noncentrality parameter for (3.6) in the system. Also the maximum likelihood (ML) estimator of β corresponds to the limited information maximum likelihood (LIML) estimator of the coefficient parameter β of an endogenous variable for (3.6) in the simultaneous equation system. We omit the details of these derivations since Anderson [4] discussed and investigated these relations in the more general framework.

From the above consideration, we obtain the next result by using arguments similar to Lemma 4.1 in Section 4.

THEOREM 3.1. *Let an estimator of β be $\hat{\beta}=\phi(\boldsymbol{H}/n, \boldsymbol{G}/((n+1)(M-1)))$, where ϕ is continuously differentiable.*

(i) *If n is fixed and $I(\beta)$ goes to infinity, or equivalently $n\ll I(\beta)$, then*

$$\mathrm{AM}_n\{\sqrt{I}(\beta)(\hat{\beta}-\beta)\}^2\geqq 1 ,$$

and the equality holds for the ML estimator.

(ii) *If both n and $I(\beta)$ go to infinity while their ratio goes to a constant, then*

$$\mathrm{AM}_n\{\sqrt{I}(\beta)(\hat{\beta}-\beta)\}^2\geqq 1+\frac{1}{\delta}$$

for $q \gg n$, and

$$\mathrm{AM}_n\{\sqrt{I}(\beta)(\hat{\beta}-\beta)\}^2 \geqq 1+\frac{1}{\delta}\left(1+\frac{\nu^2}{\delta}\right)$$

for $q=O(n)$, where $\nu^2=\lim_{q\to\infty} \lambda^2/q$. The equality holds for the ML estimator.

(iii) *If $I(\beta)$ is fixed and n goes to infinity, or alternatively $I(\beta) \ll n$, then there does not exist any consistent estimator of β.*

Some comments on this result will be helpfull. Takeuchi [18] proved (ii) assuming that Σ is known to proportional to the identity matrix when $M=1$. In the case of (i), which corresponds to the usual large sample asymptotic theory in econometrics, the ML estimator apparently attains the Crámer-Rao lower bound. However, the regular asymptotic properties of the ML estimator such as the Crámer-Rao lower bound cannot be applied to the cases of (ii) and (iii) since the number of incidental parameters increases along with the sample size N in the linear functional relationship and T in the simultaneous equation system. For the parameter sequence of (ii), which Kunitomo [11] and Morimune and Kunitomo [16] investigated assuming a known Σ with $M=1$, the ML estimator attains a possible lower bound which is larger than the information quantity. For the parameter sequence of (iii), consistent estimator cannot be constructed since the number of incidental parameters grows too fast to give enough information for estimating β. Hence the ML estimator loses even consistency in this situation. The two-stage least squares (TSLS) estimator commonly used in econometric applications, which corresponds to the least squares estimator in the linear functional relationship model (2.1)-(2.2), is inconsistent except the case (i). In the simultaneous equation system there can be large number of exogenous variables even if there are only two endogenous variables and small number (K_1) of included exogenous variables in the particular structural equation under consideration (3.6). This (large number of K_2 for each structural equation in the system) characterizes recent macro-econometric models. Kunitomo [12] discussed possible interpretations of this large K_2 asymptotic theory. If any of two endogenous variables appears in other structural equations, our interpretation in this section can be justified. We omit the details of its discussion in order to save space. Kunitomo [12] also gives a general result, which is similar to Theorem 3.1, when there are arbitrary number of endogenous and exogenous variables in the simultaneous equation system.

So far we discussed the first-order efficiency of the ML and LIML estimators in this section. Further, the third-order asymptotic optimality of the ML and LIML estimators can be proven with some minor modifications of the proof in Section 4 for alternative parameter sequences we discussed. In the more general case where the parameter of interest is a vector or matrix, which may correspond to some subsystems of the simultaneous equation system, the Edgeworth expansion of distribution of estimator becomes very complicated and our method cannot be applied to directly. However, I believe that similar results can be obtainable.

4. Proofs of theorems

Measuring all x_i, y_i and μ_i from their means, we construct the following vector: $\boldsymbol{x}'=(x_1,\cdots,x_N)\boldsymbol{P}$, $\boldsymbol{y}'=(y_1,\cdots,y_N)\boldsymbol{P}$ and $\boldsymbol{\mu}=(\mu_1,\cdots,\mu_N)\boldsymbol{P}$, where $\boldsymbol{P}=I_N-(1/N)\boldsymbol{ee}'$ and $\boldsymbol{e}'=(1,\cdots,1)$. Since there exists an $N\times N$ orthogonal matrix $\boldsymbol{R}$ such that $\boldsymbol{R}(1+\beta^2)^{1/2}\boldsymbol{\mu}/\sigma=(\lambda, 0,\cdots,0)'$ and the N-th row is $(1/N)\boldsymbol{e}'$. We define N $(=n+1)$ vectors

$$\boldsymbol{u}^*=\boldsymbol{RP}(\boldsymbol{x}+\beta\boldsymbol{y})/(\sigma(1+\beta^2)^{1/2})=(\lambda, 0,\cdots,0)'+(u_1,\cdots,u_n,0)',$$

$$\boldsymbol{v}^*=\boldsymbol{RP}(-\beta\boldsymbol{x}+\boldsymbol{y})/(\sigma(1+\beta^2)^{1/2})=(v_1,\cdots,v_n,0),$$

where $E(u_i^*)=E(v_i^*)=E(u_i^*u_j^*)=E(v_i^*v_j^*)=E(u_i^*v_j^*)=0$ $(i\neq j)$ and $E(u_i^{*2})=E(v_i^{*2})=1$ $(i, j=1,\cdots,n)$. Then vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ can be written in terms of $\boldsymbol{u}^*$ and $\boldsymbol{v}^*$ as $\boldsymbol{x}=\sigma\boldsymbol{R}(\boldsymbol{u}^*-\beta\boldsymbol{v}^*)/(1+\beta^2)^{1/2}$ and $\boldsymbol{y}=\sigma\boldsymbol{R}(\beta\boldsymbol{u}^*+\boldsymbol{v}^*)/(1+\beta^2)^{1/2}$. Defining $s_{uu}=\boldsymbol{u}^{*\prime}\boldsymbol{u}^*$, $s_{vv}=\boldsymbol{v}^{*\prime}\boldsymbol{v}^*$ and $s_{uv}=\boldsymbol{u}^{*\prime}\boldsymbol{v}^*$, we have

$$\boldsymbol{S}=\boldsymbol{Q}\begin{pmatrix} s_{uu} & s_{uv} \\ s_{uv} & s_{vv}\end{pmatrix}\boldsymbol{Q}', \qquad \boldsymbol{Q}=\sigma\begin{pmatrix} 1 & -\beta \\ \beta & 1\end{pmatrix}\Big/(1+\beta^2).$$

In the following we shall derive formal asymptotic expansions of the distribution functions of estimators in two lemmas. The discussion of validity of formal expansions is omitted to save space, but it can be obtained by following arguments given by Anderson [2] and Fujikoshi et al. [8]. The next lemma, which plays a key role in later development, is similar to Theorem 3 in Morimune and Kunitomo [16]. We note that in their derivation the asymptotic variance formula (A.8) can be rewritten as $\boldsymbol{AV}(\hat{\beta})=2((1+\beta^2)\phi_1-\beta/\rho)^2+1/\rho^2+(1+\beta^2)/\rho$, which is minimized when the first part is zero.

LEMMA 4.1. *The necessary and sufficient condition for an estimator being efficient among consistent estimators under Assumption* A *is*

$$\phi_1=-\phi_3=\frac{\beta}{\delta}, \qquad \phi_2=\frac{(1-\beta^2)}{\delta}, \tag{4.1}$$

where $\phi_i=\partial\phi/\partial h_i$ $(i=1, 2, 3)$ *evaluated at* $h_1=s_{yy}=1+\beta^2$, $h_2=s_{xy}=\beta\rho$ *and* $h_3=s_{xx}=1+\rho$.

LEMMA 4.2. *Any efficient estimator of* β *can be expressed in the* *canonical form*:

$$\sqrt{n}(\phi-\beta)=U_0+\frac{1}{\sqrt{n}}U_1+\frac{1}{n}U_2+R_1 , \tag{4.2}$$

where U_i $(i=0, 1, 2)$ *are given by* (4.4), (4.5) and (4.6), and R_1 *is a remainder term of* $o_p(n^{-1})$.

PROOF. First, we define random variables $y_{11}^*=\left(\sum_{i=2}^{N} u_i^2-n\right)\Big/\sqrt{n}$, $y_{12}^*=\sum_{i=2}^{N} u_i v_i\Big/\left(\sum_{i=2}^{N} u_i^2\right)^{1/2}$ and $y_{22}^*=\left(\sum_{i=2}^{N} v_i^2-y_{12}^2-n\right)\Big/\sqrt{n}$. Then we shall expand the distributions of estimators by Taylor's theorem in the set

$$J_n=\{|y_{ii}^*|<2\log n \ (i=1, 2),\ |y_{12}^*|<2(\log n)^{1/2},\ |u_1|<2(\log n)^{1/2} \text{ and } |v_1|<2(\log n)^{1/2}\} .$$

Then $\Pr\{J_n\}=1-O(n^{-2})$ (see Anderson [2]). A Taylor expansion of ϕ around β yields

$$\sqrt{n}(\phi-\beta)=\sum_{i=1}^{3}\phi_i h_i^*+\frac{1}{2\sqrt{n}}\sum_{i,j=1}^{3}\phi_{ij}h_i^*h_j^* +\frac{1}{6n}\sum_{i,j,k=1}^{3}\phi_{ijk}h_i^*h_j^*h_k^*+R_1 , \tag{4.3}$$

where $h_1^*=\sqrt{n}(h_1-1-\beta^2\rho)$, $h_2^*=\sqrt{1}(h_2-\beta\rho)$, $h_3^*=\sqrt{n}(h_3-1-\rho)$, and $\phi_{ij}=\partial^2\phi/\partial h_i\partial h_j$, $\phi_{ijk}=\partial^3\phi/\partial h_i\partial h_j\partial h_k$ evaluated at $h_1=1+\beta^2\rho$, $h_2=\beta\rho$, $h_3=1+\rho$, and R_1 is a polynomial of degree 3 in h_i^*, which is $O(n^{-3/2})$ and is $O((\log\sqrt{n}/\sqrt{n})^3)$ uniformly in J_n. From (4.1) and (4.3), we have

$$U_0=\frac{Z_2}{\rho}=\frac{s_{uv}}{\rho\sqrt{n}} . \tag{4.4}$$

We now differentiate (4.1) with respect to β and ρ and rearrange each term. The restrictions imposed by (4.1) are summarized as

$$\phi_{12}=-2\beta\phi_{11}+(1-\beta^2)/\delta^2 , \qquad \phi_{13}=\beta^2\phi_{11}-2\beta/\delta^2 ,$$

$$\phi_{22}=4\beta^2\phi_{11}+2\beta(\beta^2-3)/\delta^2 , \qquad \phi_{23}=-2\beta^3\phi_{11}+(5\beta^2-1)/\delta^2 ,$$

$$\phi_{33}=\beta^4+2\beta(1-\beta^2)/\delta^2 .$$

Substituting these conditions into (4.3) and rearranging each term, the second term of the right-hand-side of (4.3) becomes

$$U_1=\frac{(1+\beta^2)^2}{2}\phi_{11}Z_3^2+\frac{(1+\beta^2)}{\delta^2}\{-Z_1Z_2+Z_2Z_3-\beta Z_3^2+\beta Z_2^2\} , \tag{4.5}$$

where $Z_1=\sqrt{n}(s_{uu}/n-\delta-1)$ and $Z_3=\sqrt{n}(s_{vv}/n-1)$. We further differentiate the conditions above with respect to β and ρ, and then the resulting restrictions are summarized as:

$$\phi_{122}=-4\beta^2\phi_{111}-4\beta\phi_{112}-c_1+c_2\beta(\beta^2-3)\ ,$$

$$\phi_{123}=2\beta^3\phi_{111}+\beta^2\phi_{112}-2\beta\phi_{113}+c_1\beta+c_2(3\beta^2-1)\ ,$$

$$\phi_{133}=-\beta^4\phi_{111}+2\beta^2\phi_{113}-c_1\beta^2\phi_{11}+c_2\beta(3-\beta^2)\ ,$$

$$\phi_{222}=16\beta^3\phi_{111}+12\beta^2\phi_{112}+6c_1\beta-3c_2(\beta^4-6\beta^2+1)\ ,$$

$$\phi_{223}=-8\beta^4\phi_{111}-4\beta^3\phi_{112}+4\beta^2\phi_{113}-5c_1\beta^2+c_2\beta(9-11\beta^2)\ ,$$

$$\phi_{233}=4\beta^5\phi_{111}+\beta^4\phi_{112}-4\beta^3\phi_{113}+4c_1\beta^3+c_2(1-12\beta^2+3\beta^4)\ ,$$

$$\phi_{333}=-2\beta^6\phi_{111}+3\beta^4\phi_{113}-3c_1\beta^4+3c_2\beta(3\beta^2-1)\ ,$$

where $c_1=2\phi_{11}/\rho$ and $c_2=2/\delta^3$. Then substituting these conditions into (4.3) and rearranging each term, the third term of the right-hand-side of (4.3) becomes

$$(4.6)\qquad U_2=\frac{(1+\beta^2)}{6}\{3(\beta^2\phi_{111}+\beta\phi_{112}+\phi_{113}+2c_2\beta)Z_1Z_3^2+3(2\beta\phi_{111}+(1-\beta^2)\phi_{112}$$
$$-2\beta\phi_{113}+2c_1\beta-3c_2\beta^2)Z_2Z_3^2+((1-4\beta^2-2\beta^4)\phi_{111}-3\phi_{112}$$
$$+3\beta^2\phi_{113}-3c_1\beta^2)Z_3^3+3(-c_1+4c_2\beta)Z_2^2Z_3\}$$
$$+\frac{c_2}{2}\{-(1-\beta^2)Z_2^3+Z_1^2Z_2-2\beta Z_1Z_2^2-2Z_1Z_2Z_3\}\ .$$

Next we shall present two lemmas which give the asymptotic expansions of distributions of efficient estimators up to the order $O(n^{-1})$. We explicitly use the assumption of normality for their derivations.

LEMMA 4.3. *An asymptotic expansion of the distribution of any third-order AMDU estimator as n and λ^2 increase under Assumption* A *is given by*

$$(4.7)\qquad \Pr\left\{\frac{\hat{e}}{\tau}\leqq\xi\right\}=\Phi(\xi)-\frac{\beta\tau\xi^2}{\lambda}\phi(\xi)+\frac{\xi\phi(\xi)}{2\lambda^2\tau^2}\left\{\frac{1}{2}\left(\frac{3}{\tau^2}-1\right)-2a^2\right.$$
$$\left.+\xi^2\left(2\tau^4(\beta^2-1)+\frac{1}{2}\left(3-\frac{1}{\tau^2}\right)\right)-\beta^2\tau^4\xi^4\right\}+O(\lambda^{-3})\ ,$$

where $a=(1+\beta^2)^2(\rho\phi_{11}-c_2\beta)$.

PROOF. Let $T_0=p\lim\psi(s_{yy},s_{xy},s_{xx})=\psi(1+\beta^2,\beta\rho,1+\rho)$. Then from (4.4), the third-order asymptotic median-unbiasedness requires

$$(4.8)\qquad (1+\beta^2)^2\phi_{11}-\frac{2\beta(1+\beta^2)}{\delta^2}+T_0=0\ .$$

Differentiating this condition with respect to β and ρ, and rearranging each term, we obtain the conditions:

$$\psi_2=-2\beta\psi_1+c_2(1-\beta^4)-2c_1\beta(1+\beta^2)-(1+\beta^2)^2(2\beta\phi_{111}+\phi_{112})$$

$$\psi_3=\beta^2\psi_1-c_2(1+\beta^2)(3\beta+\beta^3)+2c_1\beta^2(1+\beta^2)+(1+\beta^2)^2(\beta^2\phi_{111}-\phi_{113}) ,$$

where $\psi_1=\partial\psi/\partial s_{yy}$, $\psi_2=\partial\psi/\partial s_{xy}$ and $\psi_3=\partial\psi/\partial s_{xx}$ evaluated at $s_{yy}=1+\beta^2\rho$, $s_{xy}=\beta\rho$ and $s_{xx}=1+\rho$. Define $T_1=\sum_{i=1}^{3}\psi_i h_i^*$. Then from (4.6),

$$\begin{aligned}(4.9)\quad T_1=&Z_1(1+\beta^2)\{-2c_2\beta-(\beta^2\phi_{111}+\beta\phi_{112}+\phi_{113})\}+Z_2(1+\beta^2)\{c_2(1+3\beta^2)\\&-(2c_1\beta+2\beta\phi_{111}+(1-\beta^2)\phi_{112}-2\beta\phi_{113})\}+Z_3(1+\beta^2)\{\psi_1-c_2\beta(1+\beta^2)\\&+\beta(2c_1\beta+\beta(\beta^2+2)\phi_{111}+\phi_{112}-\beta\phi_{113})\} .\end{aligned}$$

Let $Q_0^*=Z_2/\rho$, $Q_1^*=U_1+T_0$ and $Q_2=U_2+T_1$, where U_i are defined by (4.4), (4.5) and (4.6). Then using (4.8) and (4.9),

$$(4.10)\quad \rho^2(1+\beta^2)Q_1^*=Z_2(Z_3-Z_1)+\beta Z_2^2+\left(\frac{\phi_{11}}{\rho c_2}-\beta\right)(Z_3^2-2) ,$$

$$\begin{aligned}(4.11)\quad Q_2^*=&(1+\beta^2)\left(\frac{1}{2}Z_3^2-1\right)\{Z_1(\beta^2\phi_{111}+\beta\phi_{112}+\phi_{113}+2c_2\beta)+Z_2(2\beta\phi_{111}\\&+(1-\beta^2)\phi_{112}-2\beta\phi_{113}+2c_1\beta)-\beta Z_3(\beta(\beta^2+2)\phi_{111}+\phi_{112}\\&-\beta\phi_{113}+2c_1\beta)-3c_2\beta^2Z_2\}+Z_3^2(1+\beta^2)\left(\frac{c_1}{2}\beta^2+\frac{1}{6}(1+\beta^2)^2\phi_{111}\right)\\&+Z_3(1+\beta^2)(\psi_1-c_2\beta(1+\beta^2))+c_2(1+\beta^2)Z_2+Z_2^2Z_3(1+\beta^2)\\&\times\left(2c_2\beta-\frac{c_1}{2}\right)+\frac{c_1}{2}(1+\beta^2)(-(1-\beta^2)Z_2^3+Z_1^2Z_2\\&-2\beta Z_1Z_2^2-2Z_1Z_2Z_3) .\end{aligned}$$

Now we write $\hat{e}/\tau=\lambda(\hat{\beta}-\beta)/(1+\beta^2)=Q_0+Q_1/\lambda+Q_2/\lambda^2+R_2$, where R_2 is a remainder term of the order $o_p(n^{-1})$. By the Cornish-Fisher expansion of χ^2 random variables, we have the following canonical representation of Z_i $(i=1, 2, 3)$:

$$(4.12)\quad \begin{pmatrix}Z_1\\Z_2\\Z_3\end{pmatrix}=\begin{pmatrix}\sqrt{2}\,y_{11}+2\sqrt{\delta}\,u_1\\ y_{12}+\sqrt{\delta}\,v_1\\ \sqrt{2}\,y_{22}\end{pmatrix}+\frac{1}{\sqrt{n}}\begin{pmatrix}\frac{2}{3}\left(y_{11}^2-\frac{5}{2}\right)+u_1^2\\ \frac{1}{\sqrt{2}}y_{11}y_{12}+u_1v_1\\ \frac{2}{3}\left(y_{22}^2-4+\frac{3}{2}y_{12}^2+v_1^2\right)\end{pmatrix}+\frac{1}{n}\begin{pmatrix}\frac{1}{9\sqrt{2}}(y_{11}^3-16y_{11})\\ \frac{1}{12}(y_{11}^2-10)y_{12}\\ \frac{1}{9\sqrt{2}}(y_{22}^3-25y_{22})\end{pmatrix}+R_3 ,$$

where each component of u_1, v_1, y_{11}, y_{12} and y_{22} are mutually independent standard normal random variables. R_3 is $(1/\sqrt{n})^3$ times a polynomial of degree 4 or 2 in y_{11}, y_{12}, and y_{22} plus a remainder term, which is $o(n^{-2})$ and is $O((\log\sqrt{n}/\sqrt{n})^4)$ in J_n. Define a standard normal random variable as $W=(v_1+y_{12}/\sqrt{\delta})/\tau=E(Q_0|W)$, where $\tau=(1+1/\delta)^{1/2}$. Then W is independent of u_1, y_{11} and y_{22}, and $z=(v_1/\sqrt{\delta}-y_{12})/\tau$. Transforming v_1 and y_{12} by $(\sqrt{\delta}W+z)/(1+\delta)^{1/2}$ and $(W-\sqrt{\delta}z)/(1+\delta)^{1/2}$, we obtain

$$E(Q_1|W)=\beta\tau W^2\,, \tag{4.13}$$

$$E(Q_2|W)=\frac{W}{\tau^2}\left(\frac{5}{4}+\frac{1}{\delta}+4\tau^4\right)-\frac{1}{\tau}(1-W^2)-(1-\beta^2)\tau^2W^3\,, \tag{4.14}$$

$$E(Q_1^2|W)=\frac{1+W^2}{(1+\delta)}+\frac{\delta+W^2}{2(1+\delta)}+4\tau^4W^2+\beta^2\tau^4W^4-2\left(2+\frac{1}{\delta}\right)W^2+\frac{a^2}{\tau^2}\,, \tag{4.15}$$

where the expectation operator is taken in terms of u_1, v_1 and z in the whole space, which differs from the expectation in the set J_n by $O(\lambda^{-4})$. Finally by Fourier inversion we find (4.7).

LEMMA 4.4. *An asymptotic expansion of the distribution of any third-order AMNU estimator as n and λ^2 increase under Assumption* A *is given by*

$$\begin{aligned}\Pr\left\{\frac{\hat{e}}{\tau}\leqq\xi\right\}=\Phi(\xi)&-\frac{\beta\tau(\xi^2-1)}{\lambda}\phi(\xi)+\frac{\xi\phi(\xi)}{2\lambda^2\tau^2}\left\{\frac{1}{2}\left(\frac{3}{\tau^2}-1\right)\right.\\&-2a^2+2\tau^4(1+\beta^2+\beta^2\xi^2)-\tau^4\beta^2+\xi^2\\&\left.\times\left(2\tau^4(\beta^2-1)+\frac{1}{2}\left(3-\frac{1}{\tau^2}\right)\right)-\beta^2\tau^4\xi^4\right\}+O(\lambda^{-3})\,.\end{aligned} \tag{4.16}$$

PROOF. From (4.5), the third-order mean-unbiasedness requires

$$(1-\beta^2)^2\phi_{11}-\beta\left(\frac{(1+\beta^2)}{\delta^2}-\frac{1}{\rho}\right)+T_0=0\,. \tag{4.17}$$

Then we differentiate this condition with respect to β and ρ, and rearrange each term. The restriction implied by (4.17) is summarized as

$$\phi_2=-2\beta\phi_1+\frac{c_2}{2}(1-\beta^4)-\frac{1}{\rho^2}-2c_1\beta(1+\beta^2)-(1+\beta^2)^2(2\beta\phi_{111}+\phi_{112})$$

$$\phi_3=\beta^2\phi_1-\frac{c_2}{2}\beta(1+\beta^2)(\beta^2+3)+\frac{2\beta}{\rho^2}+2c_1\beta^2(1+\beta^2)+(1+\beta^2)^2(\beta^2\phi_{111}-\phi_{113})\,.$$

Then substitution of these conditions gives

$$T_1=Z_1(1+\beta^2)\left\{-c_2\beta+\frac{c_2}{2}\rho\beta(1+\beta^2)^2-(\beta^2\phi_{111}+\beta\phi_{112}+\phi_{113})\right\} \tag{4.18}$$

$$+Z_2(1+\beta^2)\left\{\frac{c_2}{2}(1+3\beta^2)-\frac{c_2}{2}\rho(1+\beta^2)(1+3\beta^2)-(2c_1\beta+2\beta\phi_{111}\right.$$

$$\left.+(1-\beta^2)\phi_{112}-2\beta\phi_{113})\right\}+Z_3\left\{(1+\beta^2)T_0-\frac{\beta}{\rho^3}+\frac{c_2}{2}\rho\beta(1+\beta^2)^2\right.$$

$$\left.\times(1+2\beta^2)+\beta(1+\beta^2)(2c_1\beta+\beta(\beta^2+2)\phi_{111}+\phi_{112}-\phi_{113})\right\}.$$

Let also define $Q_1^*=U_1+T_1$ and $Q_2^*=U_2+T_2$. Then by the transformation of W and z as in Lemma 4.3, we have $E(Q_0)=W$ and

$$E(Q_1|W)=\beta\tau(W^2-1)\,, \tag{4.19}$$

$$E(Q_2|W)=\frac{W}{\tau^2}\left(\frac{5}{4}+\frac{1}{\delta}\right)-\frac{1}{\tau}(1-W^2)-(1+3\beta^2)\tau^2W\,, \tag{4.20}$$

$$\begin{aligned}E(Q_1^2|W)=&\beta^2\tau^2(1-2W^2)+\frac{1}{(1+\delta)}\left(1+\delta W^2+\frac{1}{2}(\delta+W^2)\right)\\&+4\tau^4W^2+\beta^2\tau^4W^4-2\left(2+\frac{1}{\delta}\right)W^2+\frac{a^2}{\tau^2}\,,\end{aligned} \tag{4.21}$$

where the expectation operator is taken in terms of u_1, v_1, and z in the whole space, which differs the expectation in the set J_n by $O(\lambda^{-4})$. Finally, the Fourier inversion gives (4.16).

PROOFS OF THEOREMS 2.1, 2.2 AND 2.3. For the estimators $\hat{\beta}_{ML}$ and $\hat{\beta}^*$, the asymptotic expansions of their distributions are given by Kunitomo [11] and [14], respectively. Then for the ML estimator

$$\Pr\left\{\xi_1\leqq\frac{\hat{e}_1}{\tau}\leqq\xi_2\right\}-\Pr\left\{\xi_1\leqq\frac{\hat{e}_2}{\tau}\leqq\xi_2\right\}=\frac{a^2}{2\tau^2\lambda^2}\{\xi_1\phi(\xi_1)+\xi_2\phi(\xi_2)\}\geqq 0\,,$$

to terms of order n^{-1}, where $\hat{e}_1$ is the standardized $\hat{\beta}_{ML}$ (or $\hat{\beta}^*$) and $\hat{e}_2$ is any standardized third-order AMDU (or ANDU) estimator. The equality holds if and only if $a=0$.

For the angle estimator, we use $\hat{\beta}=\tan\hat{\theta}$. Putting $\xi=x+\beta\tau x^2/\lambda+\tau^2(\beta^2+1/3)x^3/\lambda^2+\cdots$, we have

$$\Pr\{\lambda(\hat{\theta}-\theta)/\tau\leqq x\}=\Phi(\xi)-\frac{\delta(1+\beta^2)}{\tau\lambda}\phi(\xi)\{\phi_{11}-2c_2\rho\beta\}+O(\lambda^{-2})\,.$$

Then the third-order asymptotic median-unbiasedness or mean-unbiasedness requires the parenthesis in the right-hand-side is zero. Hence Theorem 2.3 follows from Theorems 2.1 and 2.2.

PROOFS OF COROLLARIES. We note that

$$EL_n(\hat{\beta},\beta)=\int_0^\infty(1-\Pr\{\sqrt{n}(\hat{\beta}-\beta)\leqq y\})dh(y)$$

$$-\int_{-\infty}^{0} (\Pr\{\sqrt{n}(\hat{\beta}-\beta)\leq y\})dh(y) .$$

The proof follows immediately from this relation.

Acknowledgement

The author is grateful to Theodore W. Anderson for his encouragement and comments in the development of the thesis. The author also thanks for the comments to an earlier version of this paper from the referee and participants of the FASRAFA Conference held at Dundee University, Scotland, in August, 1983.

UNIVERSITY OF TOKYO

REFERENCES

[1] Akahira, M. and Takeuchi, K. (1981). *The Concept of Asymptotic Efficiency and Higher Order Efficiency in Statistical Estimation Theory*, No. 7, Springer-Verlag.

[2] Anderson, T. W. (1974). An asymptotic expansion of the limited information maximum likelihood estimate of a coefficient in a simultaneous equation system, *J. Amer. Statist. Ass.*, **69**, 565-572.

[3] Anderson, T. W. (1976). Estimation of linear functional relationships: approximate distributions and connections with simultaneous equations in econometrics, *J. R. Statist. Soc.*, B38, 1-38.

[4] Anderson, T. W. (1984). Estimating linear functional relationships, *Ann. Statist.*, **12**, 1-45.

[5] Anderson, T. W. and Sawa, T. (1979). Evaluation of the distribution function of the two-stage least squares estimate, *Econometrica*, **47**, 163-183.

[6] Anderson, T. W., Kunitomo, N. and Sawa, T. (1982). Evaluation of the distribution function of the limited-information maximum likelihood estimator, *Econometrica*, **50**, 1009-1027.

[7] Anderson, T. W., Kunitomo, N. and Morimune, K. (1986). Comparing single equation estimators in a simultaneous equation system, *Econom. Theory*, **2**, 1-32.

[8] Fujikoshi, Y., Morimune, K., Kunitomo, N. and Taniguchi, M. (1982). Asymptotic expansions of the distributions of the estimates of coefficients in a simultaneous equation system, *J. Econometrics*, **18**, 191-205.

[9] Fuller, W. A. (1977). Some properties of a modification of the limited information estimator, *Econometrica*, **45**, 939-953.

[10] Ghosh, J. K., Sinha, B. K. and Wieand, H. S. (1980). Second order efficiency of the MLE with respect to any bounded bowl-shaped loss function, *Ann. Statist.*, 8, 506-521.

[11] Kunitomo, N. (1980). Asymptotic expansions of the distributions of estimators in a linear functional relationship and simultaneous equations, *J. Amer. Statist. Ass.*, **75**, 693-700.

[12] Kunitomo, N. (1981). Asymptotic optimality of the limited information maximum likelihood estimator in large econometric models, *Economic Studies Quarterly*, **XXXII**, 247-266.

[13] Kunitomo, N. (1982). Asymptotic efficiency and higher order efficiency of the limited information maximum likelihood estimator in large econometric models, *Tech. Rep. No. 365*, Institute for Mathematical Studies in the Social Sciences, Stanford University.

[14] Kunitomo, N. (1986). Comparing some modified maximum likelihood estimator of a slope coefficient in a linear functional relationship, *J. Japan Statist. Soc.*, **16**, 173-185.

[15] Morimune, K. (1983). Approximate distributions of k-class estimators when the degree of overidentifiability is large compared with the sample size, *Econometrica*, **51**, 821-841.

[16] Morimune, K. and Kunitomo, N. (1980). Improving the maximum likelihood estimate in linear functional relationships for alternative parameter sequences, *J. Amer. Statist. Ass.*, **73**, 867-871.

[17] Pfanzagl, J. and Wefelmeyer, W. (1978). A third order optimum property of the maximum likelihood estimator, *J. Multivariate Anal.*, 8, 1-29.

[18] Takeuchi, K. (1972). *Contributions to the Theory of Statistical Inference in Econometrics* (in Japanese), Toyokeizai-Shinposha, Tokyo.

[19] Takeuchi, K. and Morimune, K. (1985). Asymptotic completeness of the extended maximum likelihood estimators in a simultaneous equation system, *Econometrica*, **53**, 177-200.