The Annals of Statistics
1988, Vol. 16, No. 2, 759–771

THE ASYMPTOTIC NORMAL DISTRIBUTION OF ESTIMATORS IN FACTOR ANALYSIS UNDER GENERAL CONDITIONS[1]

BY T. W. ANDERSON AND YASUO AMEMIYA

Stanford University and Iowa State University

Asymptotic properties of estimators for the confirmatory factor analysis model are discussed. The model is identified by restrictions on the elements of the factor loading matrix; the number of restrictions may exceed that required for identification. It is shown that a particular centering of the maximum likelihood estimator derived under assumed normality of observations yields an asymptotic normal distribution that is common to a wide class of distributions of the factor vectors and error vectors. In particular, the asymptotic covariance matrix of the factor loading estimator derived under the normal assumption is shown to be valid for the factor vectors containing a fixed part and a random part with any distribution having finite second moments and for the error vectors consisting of independent components with any distributions having finite second moments. Thus the asymptotic standard errors of the factor loading estimators computed by standard computer packages are valid for virtually any type of nonnormal factor analysis. The results are extended to certain structural equation models.

1. Introduction. Factor analysis is widely used in the behaviorial and social sciences, in part because of the availability of computer packages (such as LISREL) that provide estimates and their asymptotic standard errors under the assumption that the observations are normally distributed. A consequence of the results in this paper is that such asymptotic standard errors are valid for a much wider class of distributions. This asymptotic theory holds for linear functional and structural relationships, as well, if the error covariance matrix is diagonal. We shall show that it also holds for more general structural models such as the LISREL model.

The *factor analysis* model for the observable p-component random column vector $\mathbf{x}_\alpha$ can be written as

$$\mathbf{x}_\alpha = \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{f}_\alpha + \mathbf{u}_\alpha, \qquad \alpha = 1, \ldots, N, \tag{1.1}$$

where $\boldsymbol{\mu}$ is a p-component vector of parameters, $\boldsymbol{\Lambda}$ is a $p \times k$ matrix of factor loadings, $\mathbf{f}_\alpha$ is a k-component unobservable factor vector, which may contain fixed and/or random components and $\mathbf{u}_\alpha$ is a p-component unobservable random error vector. It is assumed that all $\mathbf{f}_\alpha$'s and $\mathbf{u}_\alpha$'s are uncorrelated and that $\mathscr{E}\mathbf{u}_\alpha = \mathbf{0}$ and $\mathscr{E}\mathbf{u}_\alpha\mathbf{u}'_\alpha = \boldsymbol{\Psi}$, where $\boldsymbol{\Psi} = \operatorname{diag}\{\psi_{11}, \ldots, \psi_{pp}\}$ is a $p \times p$ diagonal matrix with diagonal elements ψ_{ii}, $i = 1, \ldots, p$. The *linear functional relation*

Received November 1985; revised August 1987.

[1]This research was partly supported by NSF Grants DMS-82-19748 and DMS-86-03779.

AMS 1980 *subject classifications*. Primary 62H25; secondary 62E20.

Key words and phrases. Factor analysis, asymptotic normality, asymptotic covariance matrix, linear functional relationship, structural equation model.

model is defined by

$$\mathbf{x}_\alpha = \mathbf{z}_\alpha + \mathbf{u}_\alpha, \qquad \alpha = 1, \dots, N, \tag{1.2}$$

$$\mathbf{B}\mathbf{z}_\alpha = \boldsymbol{\beta}_0, \tag{1.3}$$

where $\mathbf{z}_\alpha$ is a fixed p-component vector, $\alpha = 1, \dots, N$, $\mathbf{B}$ is $r \times p$ and $\boldsymbol{\beta}_0$ is an r-component vector. The *linear structural relation* model is (1.2) and (1.3) with $\mathbf{z}_\alpha$ random. If $r = p - k$, (1.1) is equivalent to (1.2) and (1.3) by setting $\mathbf{z}_\alpha = \boldsymbol{\mu} + \Lambda \mathbf{f}_\alpha$ and requiring $\mathbf{B}\Lambda = \mathbf{0}$ and $\mathbf{B}\boldsymbol{\mu} = \boldsymbol{\beta}_0$.

In (1.1) Λ can be replaced by $\Lambda\mathbf{C}$ and $\mathbf{f}_\alpha$ by $\mathbf{C}^{-1}\mathbf{f}_\alpha$ to obtain an observationally equivalent model. To reduce this indeterminancy, the investigator may impose some restrictions on the parameters. If all k components of $\mathbf{f}_\alpha$ are random and if the $\mathbf{f}_\alpha$'s have a common covariance matrix Φ, then the covariance matrix of $\mathbf{x}_\alpha$ is

$$\Sigma = \Lambda\Phi\Lambda' + \Psi. \tag{1.4}$$

A traditional identification condition is that $\Phi = \mathbf{I}_k$ and $\Lambda'\Psi^{-1}\Lambda$ is diagonal. However, alternative restrictions can be placed on the elements of Λ, Φ and $\psi = (\psi_{11}, \psi_{22}, \dots, \psi_{pp})'$. In exploratory (unrestricted) factor analysis the restrictions are imposed only to remove the indeterminancy; then $\Lambda\Phi\Lambda'$ is an unrestricted positive semidefinite matrix of rank k. In confirmatory (restricted) factor analysis the investigator uses prior knowledge about the variables to formulate a hypothesis imposing restrictions on the parameters, such as certain factor loadings being 0. The model may be restricted in the sense that the number of restrictions may exceed that required for identification [Jöreskog (1969)]. A particular specification that yields identification is

$$\Lambda = \begin{pmatrix} \Lambda_1 \\ \mathbf{I}_k \end{pmatrix}. \tag{1.5}$$

In the linear functional/structural relationships there is the indeterminancy of multiplying (1.3) on the left by an arbitrary nonsingular $r \times r$ matrix. This indeterminancy can be eliminated, for example, by specifying

$$\mathbf{B} = (\mathbf{I}_r, \mathbf{B}_2). \tag{1.6}$$

If (1.5), (1.6) and $\mathbf{B}\Lambda = \mathbf{0}$ hold, then $\mathbf{B}_2 = -\Lambda_1$ and inference in the linear functional/structural relationship is identical to that in the factor analysis model. More details can be found in Anderson (1984).

A general way to parameterize exploratory and confirmatory models is to assume that restrictions are placed only on the factor loading matrix Λ and that each element of Λ can be expressed as a linear function of a $q \times 1$ parameter vector, say, $\boldsymbol{\lambda}$. Let $\operatorname{vec}\Lambda$ denote the $pk \times 1$ vector listing k columns of Λ starting from the first. Then

$$\operatorname{vec}\Lambda = \mathbf{a} + \mathbf{A}\boldsymbol{\lambda}, \tag{1.7}$$

where α is a $pk \times 1$ known vector, and $\mathbf{A}$ is a $pk \times q$ known matrix of rank q. The parameterization (1.7) covers many commonly used confirmatory factor analysis models. For example, the model identified by specified 0's and 1's in Λ satisfies (1.7), where **a** and **A** consist of 0's and 1's in specified positions. The structure (1.7) also includes cases where some factor loadings are assumed to be equal. Parameterization (1.7) with no restriction on the covariance matrix of $\mathbf{f}_\alpha$ and the error variances ψ_{ii}'s provides a unified approach to the model where the factor vector $\mathbf{f}_\alpha$ may contain fixed and random components. Furthermore, as we shall show, under this parameterization the asymptotic distribution of the estimated factor loadings is common to a very wide class of distributions of the factor vector $\mathbf{f}_\alpha$ and of the error vector $\mathbf{u}_\alpha$. We shall also show how our results based on the linear restriction (1.7) can be extended to the model where $\operatorname{vec}\Lambda$ is a nonlinear function of λ.

Under the assumption that $(\mathbf{f}_\alpha', \mathbf{u}_\alpha')'$ is normally distributed, the maximum likelihood estimators of the factor loading parameter λ, the factor covariance matrix Φ and the error variances ψ_{ii}, as well as their asymptotic standard errors, are computed by standard computer packages. We shall investigate the applicability of the asymptotic inferences based on such estimates when the $\mathbf{f}_\alpha$ and $\mathbf{u}_\alpha$ are not normally distributed and $\mathbf{f}_\alpha$ possibly contains nonstochastic components. It will be shown that the asymptotic inferences on the factor loading parameter λ in (1.7) based on the normal assumption are valid also for the model with virtually any types of $\mathbf{f}_\alpha$ and $\mathbf{u}_\alpha$ provided the p components of $\mathbf{u}_\alpha$ are independent (not just uncorrelated). We shall also indicate that our results apply not only to factor analysis but also to some more complicated structural equation models.

Statistical inference in factor analysis based on maximum likelihood was developed by Lawley (1940, 1941, 1943, 1953, 1967, 1976), Rao (1955), Anderson and Rubin (1956), Jöreskog (1967, 1969) and Jennrich and Thayer (1973). For detailed discussion of factor analysis, see, for example, Lawley and Maxwell (1971) and Anderson (1984). Properties of the normal maximum likelihood estimators under weaker assumptions were discussed by Anderson and Rubin (1956) and Amemiya, Fuller and Pantula (1987). The latter showed that for exploratory factor analysis the asymptotic distribution of the estimated factor loadings and error variances is common to a wide class of $\mathbf{f}_\alpha$ if $\mathbf{u}_\alpha$ is normally distributed. We shall extend their result to confirmatory factor analysis with the general parameterization (1.7) and to nonnormally distributed $\mathbf{u}_\alpha$. For the linear functional/structural relation model, the estimated relationship parameter $\mathbf{B}_2$ in (1.6) has been shown to have a common asymptotic distribution in a wide class of $\mathbf{z}_\alpha$ if the error $\mathbf{u}_\alpha$ is normally distributed. See, for example, Gleser (1983), Anderson (1984), Amemiya and Fuller (1984) and Chan and Mak (1985). We consider the functional/structural relation model with diagonal error covariance matrix. In terms of factor analysis, the result for $\mathbf{B}_2$ in the functional/structural relation model is equivalent to the result on Λ_1 in (1.5). Thus, our results here for Λ_1 give the results for $\mathbf{B}_2$ in the restricted functional/structural model with linearly restricted $\mathbf{B}_2$ and with nonnormally distributed error $\mathbf{u}_\alpha$.

For the factor analysis model Anderson and Rubin (1956) stated a very general theorem (Theorem 12.3) on the asymptotic distribution of the estimators obtained by maximizing the normal likelihood, which implied the validity of the asymptotic distribution for factor and error vectors with quite general distributions. However, they did not present a proof of the theorem or draw its consequences. In Section 2 we shall derive slightly more general results and obtain a modified version of Theorem 12.3 of Anderson and Rubin as a consequence. In Section 3 we shall derive several corollaries that justify the applicability of the usual large sample inference under very weak assumptions and shall discuss practical implications and extensions to some structural equation models that are more general than the factor analysis model.

2. Theorems. The estimators we consider are the maximum likelihood estimators of $\boldsymbol{\lambda}$ in (1.7), the factor covariance matrix $\boldsymbol{\Phi}$ and the error variances $\boldsymbol{\psi} = (\psi_{11}, \dots, \psi_{pp})$, derived under the assumption that $(\mathbf{f}'_\alpha, \mathbf{u}'_\alpha)'$ is normally distributed with covariance matrix block $\operatorname{diag}\{\boldsymbol{\Phi}, \boldsymbol{\Psi}\}$. Following common practice in factor analysis, we concentrate on information contained in the unbiased sample covariance matrix

$$\mathbf{S} = \frac{1}{N-1} \sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})',$$

where $\bar{\mathbf{x}} = (1/N)\Sigma_{\alpha=1}^N \mathbf{x}_\alpha$. If $\mathbf{f}_\alpha$ and $\mathbf{u}_\alpha$ are normally distributed, then $n\mathbf{S} \sim W_p(\boldsymbol{\Sigma}, n)$, where $n = N - 1$. The covariance matrix $\boldsymbol{\Sigma}$ given in (1.4) is, say, $\boldsymbol{\Sigma}(\boldsymbol{\theta})$, a function of the parameter vector $\boldsymbol{\theta} = (\boldsymbol{\lambda}', [\operatorname{vech} \boldsymbol{\Phi}]', \boldsymbol{\psi}')'$, where $\boldsymbol{\lambda}$ is given in (1.7), $\operatorname{vech} \boldsymbol{\Phi}$ is the $\frac{1}{2}k(k+1) \times 1$ vector listing the elements of $\boldsymbol{\Phi}$ that are on or below the diagonal starting with the first column, and $\boldsymbol{\psi} = (\psi_{11}, \dots, \psi_{pp})'$. Let Ω_λ, Ω_ϕ and Ω_ψ be the parameter spaces for $\boldsymbol{\lambda}$, $\operatorname{vech} \boldsymbol{\Phi}$ and $\boldsymbol{\psi}$, respectively. The set Ω_λ is in $\mathbb{R}^q$; Ω_ϕ consists of $\operatorname{vech} \boldsymbol{\Phi}$ such that $\boldsymbol{\Phi}$ is nonnegative definite; and Ω_ψ consists of $\boldsymbol{\psi}$ with nonnegative components. The Wishart likelihood based on $\mathbf{S}$ is $-n/2$ times

$$l(\boldsymbol{\theta}; \mathbf{S}) = \log|\boldsymbol{\Sigma}(\boldsymbol{\theta})| + \operatorname{tr}[\mathbf{S}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})] \tag{2.1}$$

plus terms not depending on $\boldsymbol{\theta}$. The Wishart maximum likelihood estimator $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\lambda}}', [\operatorname{vech} \hat{\boldsymbol{\Phi}}]', \hat{\boldsymbol{\psi}}')'$ is the value of $\boldsymbol{\theta}$ in $\Omega = \Omega_\lambda \times \Omega_\phi \times \Omega_\psi$ that minimizes (2.1). Note that in the computation of $\hat{\boldsymbol{\theta}}$ we allow singular estimates of $\boldsymbol{\Phi}$ (less than k factors) and zero estimates for some ψ_{ii} (Heywood cases). Later we shall assume that the true value of $\boldsymbol{\theta}$ is in the interior of Ω, that is, the true $\boldsymbol{\Phi}$ is positive definite and the true ψ_{ii} is positive.

The normal likelihood function based directly on the observations and concentrated with respect to $\hat{\mu} = \bar{x}$ is $-n/2$ times $(N/n)\log|\boldsymbol{\Sigma}(\boldsymbol{\theta})| + \operatorname{tr}[\mathbf{S}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})]$ plus terms not depending on $\boldsymbol{\theta}$. In this paper we shall treat the Wishart likelihood to avoid frequent use of the factor N/n; the estimators maximizing the normal likelihood are obtained from the Wishart estimators by replacing $\mathbf{S}$ by $(n/N)\mathbf{S}$. Of course, all of the asymptotic results hold for these estimators.

The key idea in our development of asymptotic theory is that in assessing limiting normality the estimators $\hat{\boldsymbol{\Phi}}$ and $\hat{\boldsymbol{\psi}}$ are centered around quantities depending on n. These quantities involve the unobservable sums of squares and cross products of $\mathbf{f}_\alpha$ and $\mathbf{u}_\alpha$,

$$\Phi(n) = \frac{1}{n}\sum_{\alpha=1}^{N}(\mathbf{f}_\alpha - \bar{\mathbf{f}})(\mathbf{f}_\alpha - \bar{\mathbf{f}})',$$

$$\Psi(n) = \frac{1}{n}\sum_{\alpha=1}^{N}(\mathbf{u}_\alpha - \bar{\mathbf{u}})(\mathbf{u}_\alpha - \bar{\mathbf{u}})', \tag{2.2}$$

$$\Gamma(n) = \frac{1}{n}\sum_{\alpha=1}^{N}(\mathbf{f}_\alpha - \bar{\mathbf{f}})(\mathbf{u}_\alpha - \bar{\mathbf{u}})',$$

where $\bar{\mathbf{f}} = (1/N)\Sigma_{\alpha=1}^{N}\mathbf{f}_\alpha$ and $\bar{\mathbf{u}} = (1/N)\Sigma_{\alpha=1}^{N}\mathbf{u}_\alpha$. Let $\psi(n)$ be the $p \times 1$ vector consisting of the p diagonal elements of $\Psi(n)$, and let $\psi_b(n)$ be the $\frac{1}{2}p(p-1) \times 1$ vector listing the elements of $\Psi(n)$ that are below the main diagonal. Separating the diagonal part $\psi(n)$ and the off-diagonal part $\psi_b(n)$ facilitates our development of asymptotic theory.

Our first theorem gives the consistency of the maximum Wishart likelihood estimator $\hat{\theta}$ under weak assumptions on $\Phi(n)$, $\Psi(n)$ and $\Gamma(n)$, and an identification condition. Let λ_0 be the true value of λ given in (1.7).

THEOREM 1. *In the model* (1.1) *and* (1.7) *assume*

(i) $\operatorname{plim}_{n\to\infty}\Phi(n) = \Phi_0$;
(ii) $\operatorname{plim}_{n\to\infty}\psi(n) = \psi_0$;
(iii) $\operatorname{plim}_{n\to\infty}\psi_b(n) = \mathbf{0}$;
(iv) $\operatorname{plim}_{n\to\infty}\Gamma(n) = \mathbf{0}$; *and*
(v) *for any* $\varepsilon > 0$ *there exists an* $\eta > 0$ *such that any* θ *in* Ω *with* $\|\theta - \theta_0\| > \varepsilon$ *satisfies* $\operatorname{mod}(\nu_i - 1) > \eta$ *for some* $i = 1, 2, \dots, p$, *where* $\theta_0 = [\lambda_0', \operatorname{vech}(\Phi_0)', \psi_0']'$, *the* ν_i*'s are the* p *roots of* $|\Sigma(\theta) - \nu\Sigma(\theta_0)| = 0$, *and* $\operatorname{mod}(\nu_i - 1)$ *is the absolute value of* $\nu_i - 1$.

Then

$$\operatorname*{plim}_{n\to\infty}\hat{\theta} = \theta_0.$$

PROOF. The result follows from $\operatorname{plim}_{n\to\infty}\mathbf{S} = \Sigma(\theta_0)$ and the consistency proof of Amemiya, Fuller and Pantula (1987). □

Note that if all elements of $\mathbf{f}_\alpha$ are fixed, then the probability limit in assumption (i) is the usual limit.

Let $\theta(n) = (\lambda_0', [\operatorname{vech}\Phi(n)]', \psi'(n))'$. Note that the first part of $\theta(n)$ is the true value λ_0 of λ and is free of n. Recall that $\psi(n)$ is the vector of the diagonal elements of $\Psi(n)$, and that $\Phi(n)$ and $\Psi(n)$ are defined in (2.2). The next theorem shows that the leading term in the expansion of $\hat{\theta} - \theta(n)$ is a linear function of $\psi_b(n)$ and $\Gamma(n)$ defined in (2.2).

THEOREM 2. *In the model* (1.1) *and* (1.7) *let assumptions* (i), (ii) *and* (v) *hold. Assume*

(iii-a) $\psi_b(n) = O_p(1/\sqrt{n})$;

(iv-a) $\Gamma(n) = O_p(1/\sqrt{n})$; *and*

(vi) λ_0 *is an interior point of* Ω_λ, Φ_0 *is positive definite and each element of* ψ_0 *is positive, where* Φ_0 *and* ψ_0 *are defined in* (i) *and* (ii), *respectively. The matrix*

$$\left.\frac{\partial \operatorname{vec} \Sigma(\theta)}{\partial \theta'}\right|_{\theta=\theta_0}$$

has full column rank.

Then

$$\begin{aligned}\hat{\theta} - \theta(n) &= \mathbf{C}_1(\theta_0)\psi_b(n) + \mathbf{C}_2(\theta_0)\operatorname{vec}\Gamma(n) + o_p\left(\frac{1}{\sqrt{n}}\right)\\ &= O_p\left(\frac{1}{\sqrt{n}}\right),\end{aligned}$$

where $\mathbf{C}_1(\theta_0)$ *and* $\mathbf{C}_2(\theta_0)$ *are nonstochastic matrices depending only on* $\theta_0 = [\lambda_0', (\operatorname{vech}\Phi_0)', \psi_0']'$.

PROOF. Because $\hat{\theta}$ is consistent for θ_0, an interior point of Ω, and because $l(\theta; \mathbf{S})$ in (2.1) is differentiable with respect to θ in a neighborhood of θ_0, the probability that $\hat{\theta}$ satisfies the derivative equation tends to 1 as $n \to \infty$. Thus,

$$o_p\left(\frac{1}{n}\right) = \frac{\partial l(\hat{\theta}; \mathbf{S})}{\partial \theta} = \frac{\partial l[\theta(n); \mathbf{S}]}{\partial \theta} + \frac{\partial^2 l(\theta^*; \mathbf{S})}{\partial \theta\, \partial \theta'}[\hat{\theta} - \theta(n)], \tag{2.3}$$

where θ^* is on the line segment joining $\theta(n)$ and $\hat{\theta}$. Because $\operatorname{plim}_{n\to\infty}\theta(n) = \operatorname{plim}_{n\to\infty}\hat{\theta} = \theta_0$, and because the second derivative of $l(\theta; \mathbf{S})$ with respect to θ is a continuous function of θ and $\mathbf{S}$,

$$\operatorname*{plim}_{n\to\infty} \frac{\partial^2 l(\theta^*; \mathbf{S})}{\partial \theta\, \partial \theta'} = \frac{\partial^2 l[\theta_0; \Sigma(\theta_0)]}{\partial \theta\, \partial \theta'} = \mathbf{H}_0, \tag{2.4}$$

say, where $\mathbf{H}_0$ is positive definite by assumption (vi). Also

$$\frac{\partial l[\theta(n); \mathbf{S}]}{\partial \theta} = -\mathbf{F}(n)\operatorname{vec}\{\mathbf{S} - \Sigma[\theta(n)]\}, \tag{2.5}$$

where

$$\begin{aligned}\mathbf{F}(n) &= \left\{\frac{\partial \operatorname{vec}\Sigma[\theta(n)]}{\partial \theta'}\right\}'\{\Sigma^{-1}[\theta(n)] \otimes \Sigma^{-1}[\theta(n)]\} = \mathbf{F}_0 + o_p(1),\\ \mathbf{F}_0 &= \left\{\frac{\partial \operatorname{vec}\Sigma(\theta_0)}{\partial \theta'}\right\}'\{\Sigma^{-1}(\theta_0) \otimes \Sigma^{-1}(\theta_0)\}.\end{aligned} \tag{2.6}$$

We observe that

$$\mathbf{S} - \boldsymbol{\Sigma}[\boldsymbol{\theta}(n)] = \boldsymbol{\Lambda}_0\boldsymbol{\Gamma}(n) + \boldsymbol{\Gamma}'(n)\boldsymbol{\Lambda}_0' + \boldsymbol{\Psi}(n) - \text{diag}\{\boldsymbol{\psi}(n)\}, \tag{2.7}$$

where $\text{vec}\,\boldsymbol{\Lambda}_0 = \mathbf{a} + \mathbf{A}\boldsymbol{\lambda}_0$. The diagonal elements of $\boldsymbol{\Psi}(n) - \text{diag}\{\boldsymbol{\psi}(n)\}$ are 0's, and the off-diagonal elements are elements of $\boldsymbol{\psi}_b(n)$. Hence,

$$\mathbf{S} - \boldsymbol{\Sigma}[\boldsymbol{\theta}(n)] = O_p\left(\frac{1}{\sqrt{n}}\right). \tag{2.8}$$

By (2.3)–(2.8),

$$\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(n) = \mathbf{H}_0^{-1}\mathbf{F}_0\text{vec}\{\mathbf{S} - \boldsymbol{\Sigma}[\boldsymbol{\theta}(n)]\} + o_p\left(\frac{1}{\sqrt{n}}\right) \tag{2.9}$$

and the result follows from (2.7). □

Theorem 2 shows that the term of $O_p(1/\sqrt{n})$ in the expansion of $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(n)$ depends only on the cross product $\boldsymbol{\Gamma}(n)$ and the off-diagonal part $\boldsymbol{\psi}_b(n)$ of $\boldsymbol{\Psi}(n)$. Thus, $\sqrt{n}[\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(n)]$ has a limiting distribution if the elements of $\sqrt{n}\boldsymbol{\Gamma}(n)$ and $\sqrt{n}\boldsymbol{\psi}_b(n)$ have a joint limiting distribution. The next theorem is a slight generalization of Theorem 12.3 of Anderson and Rubin (1956).

Theorem 3. *In the model* (1.1) *and* (1.7) *let assumptions* (i), (ii), (v) *and* (vi) *hold. Assume*

(iii-iv-b) *for some* $\mathbf{G}$, $\sqrt{n}([\text{vec}\,\boldsymbol{\Gamma}(n)]', \boldsymbol{\psi}_b(n))' \to_L N(\mathbf{0}, \mathbf{G})$.

Then

$$\sqrt{n}[\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(n)] \to_L N(\mathbf{0}, \mathbf{V}),$$

for some $\mathbf{V}$. *If, in addition,* $\mathbf{G}$ *depends only on* $\boldsymbol{\Phi}_0$ *and* $\boldsymbol{\psi}_0$, *then* $\mathbf{V}$ *depends only on* $\boldsymbol{\theta}_0 = [\boldsymbol{\lambda}_0', (\text{vech}\,\boldsymbol{\Phi}_0)', \boldsymbol{\psi}_0']'$.

Proof. By Theorem 2, $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(n)$ is asymptotically a linear function of $([\text{vec}\,\boldsymbol{\Gamma}(n)]', \boldsymbol{\psi}_b'(n))'$. Hence, the limiting normal distribution follows. Because $\mathbf{V}$ is a function of $\mathbf{C}_1(\boldsymbol{\theta}_0)$, $\mathbf{C}_2(\boldsymbol{\theta}_0)$ and $\mathbf{G}$, $\mathbf{V}$ is a function of $\boldsymbol{\theta}_0$ if $\mathbf{G}$ depends only on $\boldsymbol{\Phi}_0$ and $\boldsymbol{\psi}_0$. □

Note that the limiting distribution in Theorem 3 was derived under a weak set of assumptions on the $\mathbf{f}_\alpha$'s and the $\mathbf{u}_\alpha$'s. The only conditions assumed for the $\mathbf{f}_\alpha$'s and the $\mathbf{u}_\alpha$'s are assumptions (i), (ii) and (iii-iv-b). For example, there is no assumption that the $\mathbf{u}_\alpha$'s are independently and identically distributed or that the $\mathbf{f}_\alpha$'s and the $\mathbf{u}_\alpha$'s are independent. The first part of $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(n)$ is $\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0$, where $\boldsymbol{\lambda}_0$ is the true value of $\boldsymbol{\lambda}$. The use of the second part of $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(n)$, namely, $\text{vech}[\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi}(n)]$, is discussed in the next section.

In Theorem 12.3 of Anderson and Rubin (1956), instead of assumption (ii-iv-b), only the limiting normality of $\boldsymbol{\psi}_b(n)$ is assumed. However, assumption (iii-iv-b) is necessary even if we added the assumptions that the $\mathbf{f}_\alpha$'s and $\mathbf{u}_\alpha$'s are independent and that the $\mathbf{u}_\alpha$'s are independently and identically distributed. This is because the result requires the joint limiting normality of $\Gamma(n)$ and $\psi_b(n)$, not just the limiting normality of $\Gamma(n)$ and the limiting normality of $\psi_b(n)$. The point of the last assertion in the theorem is that, although the distributions of the $\mathbf{f}_\alpha$'s and $\mathbf{u}_\alpha$'s may depend on other parameters, the limiting distribution will not depend on such parameters if the limiting distribution of $\sqrt{n}([\mathrm{vec}\,\Gamma(n)]', \psi_b'(n))'$ does not.

3. Corollaries and implications. We shall show that Theorem 3 in Section 2 has important practical implications. The limiting distribution of $\sqrt{n}[\hat{\theta} - \theta(n)]$ is common to very wide classes of $\mathbf{f}_\alpha$ and $\mathbf{u}_\alpha$, and the standard asymptotic inference procedures for the factor loading parameter λ based on the normality of $(\mathbf{f}_\alpha', \mathbf{u}_\alpha')'$ are valid for virtually any type of fixed or nonnormal $\mathbf{f}_\alpha$ and of nonnormal $\mathbf{u}_\alpha$. To this end, we first present a special case of Theorem 3, where the factor vector $\mathbf{f}_\alpha$ satisfies a very weak assumption and the error vector satisfies a relatively strong assumption of normality. The normality of $\mathbf{u}_\alpha$ will be dropped in the subsequent results.

COROLLARY 1. *In the model* (1.1) *and* (1.7) *let assumptions* (v) *and* (vi) *hold. Assume*

(i-a) $\lim_{n\to\infty}\Phi(n) = \Phi_0$, *a.s.*;
(vii) *the* $\mathbf{f}_\alpha$*'s are independent of the* $\mathbf{u}_\alpha$*'s*;
(viii) *the* $\mathbf{u}_\alpha$*'s are independently and identically distributed*; *and*
(ix) $\mathbf{u}_\alpha \sim N(\mathbf{0}, \mathrm{diag}\{\psi_0\})$, *where* $\mathrm{diag}\{\psi_0\}$ *is the diagonal matrix with the p elements of* ψ_0 *on the diagonal.*

Then

$$\sqrt{n}[\hat{\theta} - \theta(n)] \to_L N(\mathbf{0}, \mathbf{V}_0), \tag{3.1}$$

for some $\mathbf{V}_0$, *where* $\mathbf{V}_0$ *depends only on* $\theta_0 = [\lambda_0', (\mathrm{vech}\,\Phi_0)', \psi_0']'$.

PROOF. To apply Theorem 3, we note that assumptions (i-a) and (viii) imply assumptions (i) and (ii), respectively. Thus, the proof will be complete when we show that under assumptions (i-a), (vii), (viii) and (ix), the assumption (iii-iv-b) in Theorem 3 holds with $\mathbf{G}$ depending only on θ_0. We outline the proof of this assertion. See Anderson and Amemiya (1985) for the detailed proof.

First, we condition on a fixed sequence $\{\mathbf{f}_\alpha\}$ satisfying $\lim_{n\to\infty}\Phi(n) = \Phi_0$. Then, conditionally,

$$\sqrt{n}([\mathrm{vec}\,\Gamma(n)]', \psi_b'(n))' \to_L N(\mathbf{0}, \mathbf{G}_0), \tag{3.2}$$

where $\mathbf{G}_0$ is a function of Φ_0 and ψ_0 only. See, for example, Lemma 1 of Amemiya and Fuller (1984) and Corollary 2.6.1 of Anderson (1971). The $\mathbf{f}_\alpha$'s and the $\mathbf{u}_\alpha$'s are independent by assumption (vii), and $\mathbf{G}_0$ depends on the given sequence $\{\mathbf{f}_\alpha\}$ only through Φ_0. Thus, it follows from assumption (i-a) and the argument used in the proofs of Theorem 2.2 of Gleser (1983) and Theorem 2.R of

Amemiya, Fuller and Pantula (1987) that (3.2) holds unconditionally. Hence, assumption (iii-iv-b) in Theorem 3 holds, and the result follows. □

The class of factor vectors $\mathbf{f}_\alpha$ satisfying the assumptions of Corollary 1 is large. If every component of $\mathbf{f}_\alpha$ is fixed, then assumption (vii) is trivially true, and the almost sure limit in assumption (i-a) reduces to the usual limit. Assumption (i-a) is satisfied for random $\mathbf{f}_\alpha$ if the $\mathbf{f}_\alpha$'s are independently and identically distributed with covariance matrix $\mathbf{\Phi}_0$. Corollary 1 shows that the covariance matrix $\mathbf{V}_0$ of the limiting normal distribution of $\sqrt{n}\,[\hat{\boldsymbol{\theta}} - \theta(n)]$ is common to the large class of $\mathbf{f}_\alpha$. A special case is the normal model where the $(\mathbf{f}_\alpha', \mathbf{u}_\alpha')$'s are independently and identically distributed according to a normal distribution with covariance matrix block $\operatorname{diag}\{\mathbf{\Phi}_0, \operatorname{diag}(\psi_0)\}$. Thus, the normal case limiting covariance matrix $\mathbf{V}_0$ is valid for a class of $\mathbf{f}_\alpha$ much wider than the normal.

In Corollary 1 the error vectors $\mathbf{u}_\alpha$'s are assumed to be normally distributed. We now show that the normality assumption on $\mathbf{u}_\alpha$ can be weakened without altering the covariance matrix $\mathbf{V}_0$ of the limiting normal distribution of $\sqrt{n}\,[\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(n)]$. The next two corollaries show that the normal case limiting covariance matrix $\mathbf{V}_0$ is valid for a large class of $\mathbf{f}_\alpha$ and a large class of $\mathbf{u}_\alpha$. Practical implications are discussed after the corollaries.

COROLLARY 2. *In the model* (1.1) *and* (1.7) *let assumptions* (i-a), (v), (vi), (vii) *and* (viii) *hold. Assume*

(ix-a) *for $i > j$ and $k > l$,*

$$E\{u_{i\alpha}u_{j\alpha}u_{k\alpha}u_{l\alpha}\} = \psi_{ii}^0\psi_{jj}^0, \qquad i = k > j = l,$$
$$= 0, \qquad \textit{otherwise},$$

where $u_{i\alpha}$ is the ith component of $\mathbf{u}_\alpha$, and ψ_{ii}^0 is the ith component of ψ_0.

Then (3.1) *holds for $\mathbf{V}_0$ defined in Corollary* 1.

PROOF. Under the assumptions the covariance matrix of $([\operatorname{vec}\Gamma(n)]', \psi_b'(n))'$ is the same as that for the case with normally distributed $\mathbf{u}_\alpha$. Thus, (3.2) in the proof of Corollary 1 holds with the same $\mathbf{G}_0$, and the result follows. □

Corollary 2 shows that the normal case limiting covariance matrix $\mathbf{V}_0$ is valid for a wide range of $\mathbf{f}_\alpha$'s and $\mathbf{u}_\alpha$'s if the off-diagonal part $\psi_b(n)$ of $\Psi(n)$ has a limiting normal distribution with covariance matrix identical to that of the case of normal $\mathbf{u}_\alpha$'s. Such an assumption on the $\mathbf{u}_\alpha$'s is not considered to be very restrictive because by the factor analysis model structure the p components of $\mathbf{u}_\alpha$ are uncorrelated. There is no restriction on the pure fourth-order moments, $Eu_{i\alpha}^4$, not even that they exist. The next corollary shows that if the p components of $\mathbf{u}_\alpha$ are independent, not just uncorrelated, then the normal case limiting covariance matrix $\mathbf{V}_0$ is valid whatever the distributions of the $\mathbf{u}_{i\alpha}$'s are. This somewhat surprising result has important practical implications that will be discussed after the corollary.

COROLLARY 3. *In the model* (1.1) *and* (1.7) *let assumptions* (i-a), (v), (vi), (vii) *and* (viii) *hold. Assume*

(ix-b) *the* $u_{i\alpha}$, $i = 1, 2, \dots, p$ *are independent.*

Then (3.1) *holds for the* $\mathbf{V}_0$ *defined in Corollary* 1.

PROOF. Assumption (ix-b) implies assumption (ix-a), and the result follows from Corollary 2. □

Thus, if we assume the independence of $u_{i\alpha}$, not just zero correlations, in the model (1.1), then the limiting distribution of $\sqrt{n}\,[\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(n)]$ is common for almost all distributions of $\mathbf{f}_\alpha$ and $\mathbf{u}_\alpha$ satisfying the model and associated second moment assumptions. Combining the results of Corollaries 2 and 3, the limiting normality and the limiting covariance matrix $\mathbf{V}_0$ for the normal $(\mathbf{f}'_\alpha, \mathbf{u}'_\alpha)'$ case are valid for a very large class of $(\mathbf{f}'_\alpha, \mathbf{u}'_\alpha)'$, if $u_{i\alpha}$, $i = 1, 2, \dots, p$, are independent, or if the mixed fourth-order moments correspond to independence. The components of the $\mathbf{f}_\alpha$'s can be either fixed or random as long as assumptions on the limit of the second moments (i-a) and on the independence from the $\mathbf{u}_\alpha$'s (vii) are satisfied. The error term $\mathbf{u}_\alpha$'s can be any independently and identically distributed random vector with independent components having finite second moments. One interpretation of the factor analysis model is that all the interdependence among the p components of the observations $\mathbf{x}_\alpha$ is explained by the factor $\mathbf{f}_\alpha$. From this point of view, the independence of the error components $u_{i\alpha}$ is a part of the model assumption, and the assumptions of Corollary 3 are satisfied by any confirmatory factor analysis model with negligible distributional restrictions on $\mathbf{f}_\alpha$ and $\mathbf{u}_\alpha$.

Standard computer packages for confirmatory factor analysis, such as LISREL, compute $\hat{\boldsymbol{\theta}}$ and print out asymptotic standard errors of the elements of $\hat{\boldsymbol{\theta}}$ obtained under the assumption of normal $\mathbf{f}_\alpha$ and normal $\mathbf{u}_\alpha$. Recall that the first part of $\sqrt{n}\,[\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(n)]$ is $\sqrt{n}\,(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0)$ and that $\mathbf{V}_0$ is the limiting covariance matrix for the normal case. Corollaries 2 and 3 show that for a large class of $(\mathbf{f}'_\alpha, \mathbf{u}'_\alpha)'$ the covariance matrix of the limiting distribution of $\sqrt{n}\,(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0)$ is the same as that for the normal $(\mathbf{f}'_\alpha, \mathbf{u}'_\alpha)'$ case. If we can assume that the measurement errors of the p variables are independent, then the standard computer packages can be used to make inferences on the loadings for virtually any type of nonnormal data.

One example of nonnormal factor analysis for which standard normal case inferences are valid is the case with heavy-tailed observations. Let the factor vector $\mathbf{f}_\alpha$ be either fixed or random, satisfying the second moment condition (i-a). Let $u_{i\alpha}$, $i = 1, 2, \dots, p$, be independent of each other and independent of $\mathbf{f}_\alpha$. Assume that for $i = 1, 2, \dots, p$, $(3/\psi_{ii})^{1/2} u_{i\alpha}$ has the Student's t distribution with 3 degrees of freedom. Then $\mathscr{E}\{u_{i\alpha}\} = 0$, $\operatorname{Var}\{u_{i\alpha}\} = \psi_{ii}$ and $\mathscr{E}\{u_{i\alpha}^3\}$ and $\mathscr{E}\{u_{i\alpha}^4\}$ do not exist, but the assumptions of Corollary 3 are satisfied and the asymptotic analysis based on the normality is valid. In this case the distribution of the observation $\mathbf{x}_\alpha$ is markedly different from normal and does not possess finite pure third and fourth moments. Another example of a nonnormal case with practical importance is a discrete factor analysis model. In many applications a standard computer package developed for the normal case is used to analyze discrete variables that may not be well approximated by normal distributions. Assume that each component of $\mathbf{f}_\alpha$ takes a few integer values, that $u_{i\alpha}$, $i = 1, 2, \dots, p$, are independent and that each $u_{i\alpha}$ takes values -1, 0 and 1. For such a discrete factor analysis model, the assumptions of Corollary 3 hold and the inferences on the factor loadings using a standard computer package are valid asymptotically.

In our general parameterization for the confirmatory and exploratory factor analyses, we assumed the linear form of the factor loading $\operatorname{vec}\Lambda = \mathbf{a} + \mathbf{A}\boldsymbol{\lambda}$. Such a linear parameterization of Λ covers practically any type of confirmatory factor analysis with restrictions on the factor loadings. However, some structural equation models that are more complicated than the factor analysis model assume that the loading matrix Λ has a nonlinear structure. Assume now the general possibly nonlinear structure $\Lambda = \Lambda(\boldsymbol{\lambda})$, where $\Lambda(\boldsymbol{\lambda})$ is twice continuously differentiable in a neighborhood of $\boldsymbol{\lambda}_0$. Then it can be shown by a straightforward extension that all of our theorems and corollaries hold with

$$\mathbf{A} = \left.\frac{\partial \operatorname{vec}\Lambda(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}'}\right|_{\boldsymbol{\lambda}=\boldsymbol{\lambda}_0}, \qquad \mathbf{a} = \operatorname{vec}\Lambda(\boldsymbol{\lambda}_0) - \mathbf{A}\boldsymbol{\lambda}_0.$$

For example, the LISREL model (structural equation model) which is used widely in the social sciences assumes that observations $(\mathbf{y}_\alpha', \mathbf{z}_\alpha')$, $\alpha = 1, 2, \dots, N$, satisfy the following:

$$\begin{aligned} &\text{measurement model for } \mathbf{y}: && \mathbf{y}_\alpha = \Lambda_y \boldsymbol{\eta}_\alpha + \boldsymbol{\varepsilon}_\alpha, \\ &\text{measurement model for } \mathbf{z}: && \mathbf{z}_\alpha = \Lambda_z \boldsymbol{\zeta}_\alpha + \boldsymbol{\delta}_\alpha, \\ &\text{structural equation model:} && \boldsymbol{\eta}_\alpha = \mathbf{B}\boldsymbol{\eta}_\alpha + \boldsymbol{\Gamma}\boldsymbol{\zeta}_\alpha, \end{aligned}$$

where $\boldsymbol{\eta}_\alpha$ and $\boldsymbol{\zeta}_\alpha$ are unobservable true values, $\boldsymbol{\varepsilon}_\alpha$ and $\boldsymbol{\delta}_\alpha$ are measurement errors; we assume the structural equation has no error. See, for example, Jöreskog and Sörbom (1981) and Everitt (1984). This LISREL model can be written as

$$\mathbf{x}_\alpha = (\mathbf{y}_\alpha', \mathbf{z}_\alpha')' = \Lambda\boldsymbol{\zeta}_\alpha + (\boldsymbol{\varepsilon}_\alpha', \boldsymbol{\delta}_\alpha')', \tag{3.3}$$

where

$$\Lambda = \begin{pmatrix} \Lambda_y (\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\Gamma} \\ \Lambda_z \end{pmatrix}.$$

Let $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ be the covariance matrices of $\boldsymbol{\zeta}_\alpha$ and $(\boldsymbol{\varepsilon}_\alpha', \boldsymbol{\delta}_\alpha')'$, respectively. Computer packages for structural equation models, such as LISREL, compute the estimators of $\mathbf{B}$, $\boldsymbol{\Gamma}$, Λ_y, Λ_z, $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ and their standard errors under the assumption of normally distributed $(\mathbf{y}_\alpha', \mathbf{z}_\alpha')$ if enough restrictions are placed on the parameters to assure the identification. If the components of the measurement error $(\boldsymbol{\varepsilon}_\alpha', \boldsymbol{\delta}_\alpha')'$ are assumed to be uncorrelated, then the LISREL model (3.3) is the factor analysis model (1.1) with nonlinearly structured Λ. Our corollaries apply to such a case. Thus, if the components of the error $(\varepsilon_\alpha', \delta_\alpha')'$ are assumed to be independent, if the covariance matrix Φ of ζ_α is unrestricted and if identification (or confirmatory) restrictions are placed on $\mathbf{B}$, Γ, Λ_y and Λ_z, then the limiting distribution of the estimator of the elements of $\mathbf{B}$, Γ, Λ_y and Λ_z is common for the factor vector ζ_α, fixed or random satisfying the assumption (i-a) and for the error $(\varepsilon_\alpha', \delta_\alpha')'$ having any distribution with finite second moments. Hence, asymptotic inferences on the structural parameters $\mathbf{B}$ and Γ and the measurement parameters Λ_y and Λ_z using the standard structural equation computer packages such as LISREL are valid for virtually any type of nonnormal LISREL models with independent measurement errors and no error in the structural equation.

It should be noted that some components of the factor vector ζ_α can be nonlinear functions of other components or of other latent variables since the identification conditions applied only to Λ do not restrict the moments of ζ_α. For example, the components can be polynomials in a latent variable.

Discussing the applications of Corollaries 2 and 3, we have concentrated on the factor loading parameter λ, because the first part of $\hat{\theta} - \theta(n)$ is $\hat{\lambda} - \lambda_0$. The remaining parts of $\hat{\theta} - \theta(n)$ correspond to $\hat{\Phi} - \Phi(n)$ and $\hat{\psi} - \psi(n)$, where $\Phi(n)$ and $\psi(n)$ are defined in (2.2). The limiting distribution of $\sqrt{n}(\hat{\theta} - \theta(n))$ does not require that the fourth-order moments of the $\mathbf{f}_\alpha$'s and $\mathbf{u}_\alpha$'s be finite, but the limiting distribution of $\sqrt{n}(\hat{\theta} - \theta_0)$ does require the existence of fourth-order moments, where $\theta_0 = [\lambda_0', (\operatorname{vech}\Phi_0)', \psi_0']'$. Thus, the limiting distribution of $\sqrt{n}[(\operatorname{vech}(\hat{\Phi} - \Phi_0))', (\hat{\psi} - \psi_0')]'$ obtained under normality will not hold for nonnormal distributions of the $\mathbf{f}_\alpha$'s and $\mathbf{u}_\alpha$'s except under restrictive conditions. However, there are practical applications of our results for the second part of $\hat{\theta} - \theta(n)$, namely, $\hat{\Phi} - \Phi(n)$. If the components of the factor vector $\mathbf{f}_\alpha$ are treated as fixed, then there is no covariance matrix Φ to be estimated, and we may be interested in making inferences on the unobservable sum of squares and cross products matrix $\Phi(n)$ of the factor vectors $\mathbf{f}_\alpha$'s for the $N = n + 1$ particular individuals in the sample. For such a fixed factor case, Corollaries 2 and 3 show that asymptotic inferences on $\Phi(n)$ assuming the normality of $\mathbf{u}_\alpha$ are valid also for a large class of nonnormal $\mathbf{u}_\alpha$ satisfying assumption (ix-a) or (ix-b).

As indicated in the Introduction, the linear functional/structural relationship when $\mathbf{B} = (\mathbf{I}, \mathbf{B}_2)$ is equivalent to the factor analysis model when $\Lambda = (\Lambda_1', \mathbf{I})'$ with $\mathbf{B}_2 = -\Lambda_1$. The matrix $\mathbf{B}_2 = -\Lambda_1$ may have other linear restrictions. In this equivalence the maximum likelihood estimator under normality of $\mathbf{B}_2$ is the negative of the estimator of Λ_1. Hence, the theorems and corollaries apply to the linear functional/structural relationship in this case. As seen in the preceding discussion, the results generalize to the case where $\mathbf{B}_2$ is a nonlinear function of an underlying parameter vector.

REFERENCES

AMEMIYA, Y. and FULLER, W. A. (1984). Estimation for the multivariate errors-in-variables model with estimated error covariance matrix. *Ann. Statist.* **12** 497–509.

AMEMIYA, Y., FULLER, W. A. and PANTULA, S. G. (1987). The asymptotic distributions of some estimators for a factor analysis model. *J. Multivariate Anal.* **22** 51–64.

ANDERSON, T. W. (1971). *The Statistical Analysis of Time Series*. Wiley, New York.

ANDERSON, T. W. (1984). Estimating linear statistical relationships. *Ann. Statist.* **12** 1–45.

ANDERSON, T. W. and AMEMIYA, Y. (1985). The asymptotic normal distribution of estimators in factor analysis under general conditions. Technical Report no. 12, Econometric Workshop, Stanford Univ.

ANDERSON, T. W. and RUBIN, H. (1956). Statistical inference in factor analysis. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **5** 111–150. Univ. California Press.

CHAN, N. N. and MAK, T. K. (1985). Estimation in multivariate errors-in-variables models. *Linear Algebra Appl.* **70** 197–207.

EVERITT, B. S. (1984). *An Introduction to Latent Variable Models*. Chapman and Hall, New York.

GLESER, L. J. (1983). Functional, structural and ultrastructural errors-in-variables models. *Proc. Bus. Econ. Statist. Sec.* 57–66. Amer. Statist. Assoc., Washington.

JENNRICH, R. I. and THAYER, D. T. (1973). A note on Lawley's formulas for standard errors in maximum likelihood factor analysis. *Psychometrika* **38** 571–580.

JÖRESKOG, K. G. (1967). Some contributions to maximum likelihood factor analysis. *Psychometrika* **32** 443–482.

JÖRESKOG, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika* **34** 183–202.

JÖRESKOG, K. G. and SÖRBOM, D. (1981). *LISREL V User's Guide*. International Educational Services, Chicago.

LAWLEY, D. N. (1940). The estimation of factor loadings by the method of maximum likelihood. *Proc. Roy. Soc. Edinburgh Sect. A* **60** 64–82.

LAWLEY, D. N. (1941). Further investigations in factor estimation. *Proc. Roy. Soc. Edinburgh Sect. A* **61** 176–185.

LAWLEY, D. N. (1943). The application of the maximum likelihood method to factor analysis. *British J. Psych.* **33** 172–175.

LAWLEY, D. N. (1953). A modified method of estimation in factor analysis and some large sample results. In *Uppsala Symposium on Psychological Factor Analysis, 17–19 March 1953* 34–42. Almqvist and Wicksell, Uppsala.

LAWLEY, D. N. (1967). Some new results in maximum likelihood factor analysis. *Proc. Roy. Soc. Edinburgh Sect. A* **67** 256–264.

LAWLEY, D. N. (1976). The inversion of an augmented information matrix occurring in factor analysis. *Proc. Roy. Soc. Edinburgh Sect. A* **75** 171–178.

LAWLEY, D. N. and MAXWELL, A. E. (1971). *Factor Analysis as a Statistical Method*. Elsevier, New York.

RAO, C. R. (1955). Estimation and tests of significance in factor analysis. *Psychometrika* **20** 93–111.

DEPARTMENT OF STATISTICS
SEQUOIA HALL
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305

DEPARTMENT OF STATISTICS
SNEDECOR HALL
IOWA STATE UNIVERSITY
AMES, IOWA 50011