The Annals of Statistics
1980, Vol. 8, No. 2, 407–422

PROPERTIES OF SOME ESTIMATORS FOR THE ERRORS-IN-VARIABLES MODEL[1]

BY WAYNE A. FULLER

Iowa State University

The limiting behavior of estimators for several errors-in-variables models is investigated. It is assumed that an estimator of the covariance matrix of the measurement error is available. Models are delineated on the basis of the prior knowledge of the error structure. In all cases the limiting distribution of the estimators, standardized by $n^{\frac{1}{2}}$, is normal. Modifications of the estimators that guarantee finite moments and improve the small sample behavior of the estimators are presented.

1. Introduction. The statistical model containing errors of measurement dates from the 19th century. Reviews of the literature and discussions of the problem are contained in Moran (1971), Madansky (1959), Kendall and Stuart (1961), Cochran (1968), and Malinvaud (1970). Koopmans (1937) derived the maximum likelihood estimator for the case of known error covariance matrix. He also presented the approximate covariance matrix of the estimated coefficients under the assumption that the variance of the measurement errors was small relative to variation in the true values of the variables. Following the work of Lawley (1953), Malinvaud (1970) presented an expression for the covariance matrix of the limiting distribution of the maximum likelihood estimator.

Dorff and Gurland (1961) investigated the variance of a number of estimators for the model with a single independent variable. In particular they demonstrated that the maximum likelihood estimator using estimated error variances was superior to other estimators of the moment type. Lord (1960) and DeGracie and Fuller (1972) presented estimators for this model with application to the analysis of covariance. Robertson (1974) presented approximations for several models with a single independent variable. Schneeweiss (1976) obtained the limiting distribution of the estimator for the model with the covariance matrix of the measurement error in the independent variables known.

Anderson (1951) studied maximum likelihood estimation of the errors-in-variables model placing the problem in the context of estimating regression parameters subject to linear restrictions. In this model the number of regression coefficients is fixed. Limiting properties of the estimators were obtained under the assumption that the variance of the coefficients converged to zero. The simultaneous equation model of econometrics is an example of the problem. See Anderson and Rubin

Received April 1972; revised September 1978.

[1]This work was partly supported by Joint Statistical Agreement 73-2 with the U.S. Bureau of the Census.

AMS 1970 *subject classifications.* Primary 62J99, 62H25; secondary 62P15, 62P20, 62P25.

Key words and phrases. Errors in variables, response error, functional model.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics*.
www.jstor.org

(1949), (1950) and Anderson (1976). Villegas (1961), (1966) also studied parameter estimation for the errors-in-variables model wherein the number of points at which observations are made is fixed.

We investigate the large sample properties of the estimators of the errors-in-variables model when the covariance matrix of the measurement error is estimated. We assume that the covariance matrix of the measurement error is fixed (not decreasing with sample size) and assume that the variance of the estimator of the error covariance matrix is inversely proportional to the sample size. Estimators associated with models containing different amounts of information on the error structure are presented.

Most of the estimators presently in the literature do not possess finite moments. We modify these estimators so that they possess finite moments, leaving the limiting distribution of the estimators unchanged.

2. The model. Let $\{\mathbf{x}_{t.} : t = 1, 2, \cdots\}$ denote a fixed sequence of k dimensional row vectors. Let

$$\begin{aligned} \mathbf{y} &= \mathbf{x}\boldsymbol{\beta}_1, \\ \mathbf{X} &= \mathbf{x} + \mathbf{u}, \\ \mathbf{Y} &= \mathbf{y} + \mathbf{e}, \end{aligned} \tag{2.0}$$

where $\mathbf{x}$ is an $n \times k$ matrix whose tth row is $\mathbf{x}_{t.}$; $\boldsymbol{\beta}_1$ is a vector of k unknown parameters; $\boldsymbol{\varepsilon} = (\mathbf{e} : \mathbf{u})$ is an $n \times (k + 1)$ matrix of random variables whose rows are independently and identically distributed as a multivariate normal random variable with mean zero and covariance matrix $\boldsymbol{\Sigma}$. That is, $\boldsymbol{\varepsilon}_t = (e_t : \mathbf{u}_t) \sim \text{NID}(\mathbf{0}, \boldsymbol{\Sigma})$, for $t = 1, 2, \cdots$, where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_e^2 & \boldsymbol{\Sigma}_{eu} \\ \boldsymbol{\Sigma}_{ue} & \boldsymbol{\Sigma}_{uu} \end{pmatrix}.$$

It is assumed that

$$\sigma_e^2 - 2\boldsymbol{\Sigma}_{eu}\boldsymbol{\beta}_1 + \boldsymbol{\beta}_1'\boldsymbol{\Sigma}_{uu}\boldsymbol{\beta}_1 > 0.$$

We shall sometimes condense the notation letting $\mathbf{Z} = (\mathbf{Y} : \mathbf{X})$ and $\mathbf{z} = (\mathbf{y} : \mathbf{x})$. In this notation the model is given by

$$\begin{aligned} \mathbf{z}\boldsymbol{\beta} &= 0, \\ \mathbf{Z} &= \mathbf{z} + \boldsymbol{\varepsilon}, \end{aligned} \tag{2.1}$$

where $\boldsymbol{\beta}' = (1, -\boldsymbol{\beta}_1')$. The matrices $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$, $\mathbf{u}$, $\mathbf{e}$, and $\boldsymbol{\varepsilon}$ could be subscripted with n, but to simplify the notation we have omitted the subscript.

We shall investigate the limiting behavior of estimators of $\boldsymbol{\beta}$ as $n \to \infty$. We also assume: the matrix $n^{-1}\mathbf{x}'\mathbf{x}$ is positive definite for all $n > k$ and

$$\lim_{n\to\infty} n^{-1}\mathbf{x}'\mathbf{x} = \overline{\mathbf{M}}_{xx}, \tag{2.2}$$

where $\overline{\mathbf{M}}_{xx}$ is a positive definite matrix.

We identify five cases on the basis of the amount of information available on $\boldsymbol{\Sigma}$.

CASE (i). It is assumed that $\mathbf{e}$ is the sum of two components, $\mathbf{e} = \mathbf{w} + \mathbf{q}$, where $\mathbf{w}$ is a vector of errors of measurement and $\mathbf{q}$ is a vector of errors in the equation. It is assumed that $(\mathbf{w}, \mathbf{u})$ is independent of $\mathbf{q}$ and that the elements of $\mathbf{q}$ are normal independent $(0, \sigma_q^2)$ random variables, $\sigma_q^2 > 0$. Let

$$\mathbf{S} = \begin{pmatrix} S_{ww} & \mathbf{S}_{wu} \\ \mathbf{S}_{uw} & \mathbf{S}_{uu} \end{pmatrix}$$

denote an estimator of

$$E\{(w_t, \mathbf{u}_t)(w_t, \mathbf{u}_t)'\} = \begin{pmatrix} \sigma_w^2 & \boldsymbol{\Sigma}_{wu} \\ \boldsymbol{\Sigma}_{uw} & \boldsymbol{\Sigma}_{uu} \end{pmatrix}$$

that is distributed as a multiple of a Wishart matrix with d degrees of freedom. The estimator $\mathbf{S}$ is assumed to be independent of $(\mathbf{u}, \mathbf{w}, \mathbf{q})$. Because $\mathbf{q}$ and $\mathbf{u}$ are independent, $\mathbf{S}_{wu}$ is also an estimator of $\boldsymbol{\Sigma}_{eu}$ and, in situations where no confusion will result, we may denote $\mathbf{S}_{wu}$ by $\mathbf{S}_{eu}$.

The variance of q_t is assumed unknown. Thus we have available an estimator of the covariance structure of the measurement error, but the variance of the error in the equation σ_q^2 is unknown. Malinvaud (1970, page 374) discusses the distinction between errors of measurement (errors in the variables) and errors in the equation.

The degrees of freedom d are assumed to satisfy $d = \eta^{-1}n$ where η is a fixed positive number. To assume that

$$\lim_{n\to\infty} n^{-1}d = \eta^{-1}$$

would be sufficient, but to simplify the presentation we use the assumption $d = \eta^{-1}n$. As before, we have suppressed the functional dependence of S_{ww}, $\mathbf{S}_{wu}$, $\mathbf{S}_{uu}$ and d upon n.

An example of Case (i) is the estimation of an income elasticity from survey data subject to response error. Battese, et al. (1976) give examples of the estimation of matrices such as $\boldsymbol{\Sigma}_{uu}$ and $\boldsymbol{\Sigma}_{uw}$.

CASE (ii). It is assumed that the covariance between $\mathbf{u}_t$ and e_t is known to be zero ($\boldsymbol{\Sigma}_{ue} = \mathbf{0}$), and that an estimator $\mathbf{S}_{uu}$ of $\boldsymbol{\Sigma}_{uu}$ satisfying the assumptions of Case (i) is available. That is, it is assumed that $E\{\mathbf{S}_{uu}\} = \boldsymbol{\Sigma}_{uu}$ and that a multiple of $\mathbf{S}_{uu}$ has a Wishart distribution with d degrees of freedom, where $d = \eta^{-1}n$. It is assumed that $\sigma_e^2 > 0$ is unknown or that $\mathbf{e} = \mathbf{q} + \mathbf{w}$ and $\sigma_q^2 > 0$ is unknown.

CASE (iia). The covariance matrix $\boldsymbol{\Sigma}$ is known to be diagonal. Unbiased estimates of the diagonal elements of $\boldsymbol{\Sigma}_{uu}$ are available, where the estimates are multiples of independent chi-square random variables. This situation arises when it is known that the measurement errors for the elements of $\mathbf{x}_{t.}$ are independent. The model is used frequently when the elements of $\mathbf{x}_{t.}$ are scores on psychological instruments. See Warren et al. (1974) for an application in sociology.

It is assumed that the estimates of the variances of the errors of measurement for the variables $X_1, X_2, \cdots, X_k$ are based upon degrees of freedom, $d_1, d_2, \cdots, d_k$, respectively, where $d_1 = \eta_1^{-1}n$, $d_2 = \eta_2^{-1}n, \cdots, d_k = \eta_k^{-1}n$ and $\eta_1, \eta_2, \cdots, \eta_k$ are fixed positive numbers. As in Case (ii) it is assumed that $\sigma_e^2 > 0$ is unknown or that $\mathbf{e} = \mathbf{q} + \mathbf{w}$ and $\sigma_q^2 > 0$ is unknown.

CASE (iii). An estimator, $\tilde{\mathbf{S}}$, of a multiple of $\mathbf{\Sigma}$ satisfying the assumptions of Case (i) is available. Note that Case (iii) differs from Cases (i), (ii) and (iia) in that an estimator of σ_e^2 is available. Case (iii) is a generalization of the classical errors-in-variables model wherein it is assumed that $\mathbf{\Sigma}$ is known up to a multiple. An experiment where $\eta^{-1} + 1$ observations on the tth row of $\mathbf{Z}$, $\mathbf{Z}_{t.}$, are available for each $\mathbf{z}_{t.}$, $t = 1, 2, \cdots, n$ is an example of Case (iii). In such an experimental situation the multiple is known to be one, i.e., $E\{\tilde{\mathbf{S}}\} = \mathbf{\Sigma}$. An example of Case (iii) is contained in Fuller (1978).

CASE (iiia). The matrix $\mathbf{\Sigma}$ is known to be diagonal. Unbiased estimators of $\sigma_e^2, \sigma_{u(1)}^2, \sigma_{u(2)}^2, \cdots, \sigma_{u(k)}^2$ are available, where the estimators are multiples of independent chi-square random variables and $\sigma_{u(i)}^2$ is the variance of the measurement error in X_i. It is assumed that the estimators of the variances of the errors of measurement for $Y, X_1, X_2, \cdots, X_k$ are based upon degrees of freedom $d_0, d_1, d_2, \cdots, d_k$, respectively, where $d_0 = \eta_0^{-1}n$, $d_1 = \eta_1^{-1}n$, $d_2 = \eta_2^{-1}n, \cdots, d_k = \eta_k^{-1}n$ and $\eta_0, \eta_1, \eta_2, \cdots, \eta_k$ are fixed positive numbers. This model occurs in psychology and sociology.

3. Estimation for case (i).

We consider the estimator

$$\tilde{\boldsymbol{\beta}}_1 = (\hat{\mathbf{H}} + \alpha n^{-1}\mathbf{S}_{uu})^{-1}(\hat{\mathbf{N}} + \alpha n^{-1}\mathbf{S}_{uw}), \tag{3.1}$$

where $\alpha > 0$ is a fixed real number,

$$\begin{aligned} \hat{\mathbf{H}} &= \hat{\mathbf{M}}_{XX} - \mathbf{S}_{uu} && \text{if } \hat{\gamma} \geq 1 + n^{-1} \\ &= \hat{\mathbf{M}}_{XX} - (\hat{\gamma} - n^{-1})\mathbf{S}_{uu}, && \text{if } \hat{\gamma} < 1 + n^{-1}, \\ \hat{\mathbf{N}} &= \hat{\mathbf{M}}_{XY} - \mathbf{S}_{uw}, && \text{if } \hat{\gamma} \geq 1 + n^{-1} \\ &= \hat{\mathbf{M}}_{XY} - (\hat{\gamma} - n^{-1})\mathbf{S}_{uw}, && \text{if } \hat{\gamma} < 1 + n^{-1}, \end{aligned}$$

$$\hat{\mathbf{M}} = \begin{pmatrix} \hat{M}_{YY} & \hat{\mathbf{M}}_{YX} \\ \hat{\mathbf{M}}_{XY} & \hat{\mathbf{M}}_{XX} \end{pmatrix} = \frac{1}{n}\begin{pmatrix} \mathbf{Y}'\mathbf{Y} & \mathbf{Y}'\mathbf{X} \\ \mathbf{X}'\mathbf{Y} & \mathbf{X}'\mathbf{X} \end{pmatrix} = \frac{1}{n}\mathbf{Z}'\mathbf{Z},$$

and $\hat{\gamma}$ is the smallest root of $|\hat{\mathbf{M}} - \gamma\mathbf{S}| = 0$. Observe that $\hat{\mathbf{H}}$ is an estimator of the mean squares and products of the true values, say $\mathbf{M}_{xx}$, where

$$\mathbf{M}_{xx} = n^{-1}\mathbf{x}'\mathbf{x}.$$

In a similar manner $\hat{\mathbf{N}}$ is an estimator of

$$\mathbf{M}_{xy} = n^{-1}\mathbf{x}'\mathbf{y}.$$

The modification associated with the computation of $\hat{\gamma}$ guarantees that $\hat{\mathbf{H}}$ is a positive definite matrix, that the estimator of σ_q^2 is positive and that the estimator of $\boldsymbol{\beta}_1$ possesses finite variance. If $\hat{\gamma}$ is close to one it is advisable to compute $\hat{l}$ where $\hat{l}$

is the smallest root of

$$|\hat{\mathbf{M}}_{XX} - l\mathbf{S}_{uu}| = 0,$$

as a model check. By Theorem 5, $(n - k + 1)^{-1}n\hat{l}$ is approximately distributed as Snedecor's F with $n - k + 1$ and d degrees of freedom when the rank of $\mathbf{M}_{xx}$ is $k - 1$. Thus, in practice, a small $\hat{l}$ may call into question the assumption that $\mathbf{M}_{xx}$ is nonsingular.

The modification associated with α was suggested by a study of the moments through terms of order n^{-2}. The α-modification gives an estimator that is similar to the "k-class" estimators used in simultaneous equation estimation, see, for example, Johnston (1972, page 388). Our situation differs from that in simultaneous equations, however. In our model $\mathbf{S}_{uu}$ is $\mathcal{O}_p(1)$ while the analogous statistic in the simultaneous equation problem is $\mathcal{O}_p(n^{-1})$.

THEOREM 1. *Let the assumptions of Case* (i) *of model* (2.0) *hold. Then the limiting distribution of* $n^{\frac{1}{2}}(\tilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1)$ *is normal with mean zero and covariance matrix*

$$\overline{\mathbf{M}}_{xx}^{-1}\sigma_v^2 + \overline{\mathbf{M}}_{xx}^{-1}\left[(\sigma_v^2 + \eta\sigma_r^2)\boldsymbol{\Sigma}_{uu} + (1 + \eta)\boldsymbol{\Sigma}_{uv}\boldsymbol{\Sigma}_{vu}\right]\overline{\mathbf{M}}_{xx}^{-1},$$

where $\mathbf{v} = \mathbf{e} - \mathbf{u}\boldsymbol{\beta}_1$, $\mathbf{r} = \mathbf{w} - \mathbf{u}\boldsymbol{\beta}_1$, $\boldsymbol{\Sigma}_{ur} = \boldsymbol{\Sigma}_{uv} = \boldsymbol{\Sigma}_{ue} - \boldsymbol{\Sigma}_{uu}\boldsymbol{\beta}_1$, $\sigma_v^2 = \sigma_e^2 - 2\boldsymbol{\Sigma}_{eu}\boldsymbol{\beta}_1 + \boldsymbol{\beta}_1'\boldsymbol{\Sigma}_{uu}\boldsymbol{\beta}_1$, *and* $\sigma_r^2 = \sigma_w^2 - 2\boldsymbol{\Sigma}_{eu}\boldsymbol{\beta}_1 + \boldsymbol{\beta}_1'\boldsymbol{\Sigma}_{uu}\boldsymbol{\beta}_1$.

PROOF. We note that the correction associated with $\hat{\gamma}$ occurs with probability that is $\mathcal{O}(n^{-2})$. Hence, we can write

$$\begin{aligned}\tilde{\boldsymbol{\beta}}_1 &= \left[\hat{\mathbf{M}}_{XX} - (1 - \alpha n^{-1})\mathbf{S}_{uu}\right]^{-1}\left[\hat{\mathbf{M}}_{XY} - (1 - \alpha n^{-1})\mathbf{S}_{uw}\right] + \mathcal{O}_p(n^{-2})\\ &= \left[\hat{\mathbf{M}}_{XX} - (1 - \alpha n^{-1})\mathbf{S}_{uu}\right]^{-1}\left[\hat{\mathbf{M}}_{XX}\boldsymbol{\beta}_1 + \hat{\mathbf{M}}_{Xv} - (1 - \alpha n^{-1})\mathbf{S}_{uu}\boldsymbol{\beta}_1\right.\\ &\quad \left.- (1 - \alpha n^{-1})\mathbf{S}_{ur}\right] + \mathcal{O}_p(n^{-2}),\end{aligned}$$

where

$$\begin{aligned}\hat{\mathbf{M}}_{Xv} &= n^{-1}\mathbf{X}'\mathbf{v},\\ \mathbf{S}_{ur} &= \mathbf{S}_{uw} - \mathbf{S}_{uu}\boldsymbol{\beta}_1,\\ \mathbf{v} &= \mathbf{Y} - \mathbf{X}\boldsymbol{\beta}_1.\end{aligned}$$

It follows that

$$\tilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1 = \left[\hat{\mathbf{M}}_{XX} - \mathbf{S}_{uu}\right]^{-1}\left[\hat{\mathbf{M}}_{Xv} - \mathbf{S}_{ur}\right] + \mathcal{O}_p(n^{-1}).$$

The matrices $\hat{\mathbf{M}}_{XX}$, $\hat{\mathbf{M}}_{XY}$, $\mathbf{S}_{uu}$, and $\mathbf{S}_{uw}$ are all of the sample moment type. Therefore

$$\begin{aligned}\hat{\mathbf{M}}_{XX} - \mathbf{S}_{uu} &= \mathbf{M}_{xx} + \mathcal{O}_p(n^{-\frac{1}{2}}),\\ \hat{\mathbf{M}}_{Xv} - \mathbf{S}_{ur} &= \mathcal{O}_p(n^{-\frac{1}{2}}),\end{aligned}$$

and

$$\tilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1 = \mathbf{M}_{xx}^{-1}\left[\hat{\mathbf{M}}_{Xv} - \mathbf{S}_{ur}\right] + \mathcal{O}_p(n^{-1})$$
$$= \mathbf{M}_{xx}^{-1}\left[n^{-1}\mathbf{x}'\mathbf{v} + n^{-1}\mathbf{u}'\mathbf{v} - \mathbf{S}_{ur}\right] + \mathcal{O}_p(n^{-1}).$$

Consider the linear combination

$$n^{\frac{1}{2}}\boldsymbol{\lambda}'\mathbf{M}_{xx}(\tilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1) = n^{-\frac{1}{2}}\Sigma_{t=1}^{n}\Sigma_{i=1}^{k}\lambda_i\{x_{ti}v_t + (u_{ti}v_t - \sigma_{u_i v})\}$$
$$-\eta^{\frac{1}{2}}d^{\frac{1}{2}}\Sigma_{i=1}^{k}\lambda_i(s_{u(i)r} - \sigma_{u(i)r}) + \mathcal{O}_p(n^{-\frac{1}{2}})$$
$$= S_{1n} + S_{2n} + \mathcal{O}_p(n^{-\frac{1}{2}}), \text{ say,}$$

where $\boldsymbol{\lambda}$ is an arbitrary real vector and $\sigma_{u(i)v} = \sigma_{u(i)r}$. The random variables

$$g_t = \Sigma_{i=1}^{k}\lambda_i\{x_{ti}v_t + (u_{ti}v_t - \sigma_{u(i)v})\}$$
$$= C_t v_t + h_t v_t - \sigma_{hv},$$

where

$$C_t = \Sigma_{i=1}^{k}\lambda_i x_{ti}$$
$$h_t = \Sigma_{i=1}^{k}\lambda_i u_{ti}$$

are independent with zero mean. Because $(e_t, \mathbf{u}_t)$ is normal,

$$\operatorname{Var}\{g_t\} = C_t^2\sigma_v^2 + \sigma_h^2\sigma_v^2 + \sigma_{hv}^2.$$

Letting

$$V_n = \Sigma_{t=1}^{n}\operatorname{Var}\{g_t\},$$

we have

$$V_n^{-1}\Sigma_{t=1}^{n}\int_{R_{1n}}(C_t v + hv - \sigma_{hv})^2 dF(h, v) \tag{3.2}$$
$$\leqslant V_n^{-1}\Sigma_{t=1}^{n}a_t^2\int_{R_{2n}}(|v| + |hv - \sigma_{hv}|)^2 dF(h, v),$$

where $F(h, v)$ is the distribution function of the random vector (h, v),

$$R_{1n} = \{(h, v) : (C_t v + hv - \sigma_{hv})^2 > \varepsilon^2 V_n\},$$
$$R_{2n} = \{(h, v) : (|v| + |hv - \sigma_{hv}|)^2 > \varepsilon^2 V_n[\sup_{1\leqslant t\leqslant n}a_t^2]^{-1}\}$$

and

$$a_t^2 = \max(C_t^2, 1), \qquad \sigma_h^2 > 0$$
$$= C_t^2, \qquad \text{otherwise.}$$

Because $[\operatorname{Var}\{g_t\}]^{-1}a_t^2$ is bounded, the ratio

$$V_n^{-1}\Sigma_{t=1}^{n}a_t^2 \tag{3.3}$$

is bounded. By assumption (2.2)

$$\lim_{n\to\infty}\left(V_n^{-1}\sup_{1\leqslant t\leqslant n}a_t^2\right)^{-1} = \infty,$$

and, as (w, v) is bivariate normal,

$$\lim_{n\to\infty}\int_{R_{2n}}(|v| + |hv - \sigma_{hv}|)^2 dF(h, v) = 0. \tag{3.4}$$

It follows from (3.2), (3.3), and (3.4) that the Lindeberg condition is satisfied for S_{1n}.

The quantity S_{2n} is independent of S_{1n} for all n, and S_{2n} converges to a normal random variable. Hence, the sum $S_{1n} + S_{2n}$ converges to a normal random variable and the k-dimensional vector $n^{\frac{1}{2}}(\tilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1)$ converges in distribution to a k-dimensional normal random variable. Now

$$E\{[n^{-1}\mathbf{x}'\mathbf{v} + n^{-1}\mathbf{u}'\mathbf{v} - \mathbf{S}_{ur}][n^{-1}\mathbf{v}'\mathbf{x} + n^{-1}\mathbf{v}'\mathbf{u} - \mathbf{S}_{ru}]\}$$
$$= n^{-1}\mathbf{M}_{xx}\sigma_v^2 + n^{-1}(\boldsymbol{\Sigma}_{uu}\sigma_v^2 + \boldsymbol{\Sigma}_{uv}\boldsymbol{\Sigma}_{vu}) + d^{-1}(\boldsymbol{\Sigma}_{uu}\sigma_r^2 + \boldsymbol{\Sigma}_{ur}\boldsymbol{\Sigma}_{ru})$$

and the result is established. □

By setting $\eta = 0$ in the conclusion of Theorem 1, one obtains the variance of the limiting distribution of $n^{\frac{1}{2}}(\tilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1)$ for known $\boldsymbol{\Sigma}_{uu}$ and $\boldsymbol{\Sigma}_{ue}$.

It is demonstrated in the Appendix that, for $\{|\mathbf{x}_{t.}|\}$ uniformly bounded,

$$E\{\tilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1\} = -\mathbf{M}_{xx}^{-1}\{n^{-1}(k + 1 - \alpha) + (n^{-1} + d^{-1})([\mathrm{tr}(\boldsymbol{\Sigma}_{uu}\mathbf{M}_{xx}^{-1})]\mathbf{I}$$
$$+\boldsymbol{\Sigma}_{uu}\mathbf{M}_{xx}^{-1})\}\boldsymbol{\Sigma}_{uv} + \mathcal{O}(n^{-2})$$

and that, to order n^{-2}, the mean square error of $\tilde{\boldsymbol{\beta}}_1$ is smaller for $\alpha = k + 4 + 2\eta$ than for any smaller α.

Of some interest is the hypothesis that ordinary least squares yields an unbiased estimator of $\boldsymbol{\beta}_1$. The observations $\mathbf{Y}$ may be expressed as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta}_1 + \mathbf{v},$$

where $\mathbf{v} = \mathbf{e} - \mathbf{u}\boldsymbol{\beta}_1$. Assuming $\boldsymbol{\Sigma}_{uu}$ to be nonsingular, the least squares estimator of $\boldsymbol{\beta}_1$ will be unbiased if $E\{\mathbf{v}|\mathbf{X}\} = E\{\mathbf{v}|(\mathbf{x} + \mathbf{u})\} = \mathbf{u}\boldsymbol{\Sigma}_{uu}^{-1}\boldsymbol{\Sigma}_{uv} = \mathbf{0}$. In turn, this conditional expectation will be zero if $\boldsymbol{\Sigma}_{uv} = E\{\mathbf{u}_t'v_t\} = \boldsymbol{\Sigma}_{ue} - \boldsymbol{\Sigma}_{uu}\boldsymbol{\beta}_1 = \mathbf{0}$. Thus, the ordinary least squares estimator will be unbiased when $\boldsymbol{\beta}_1 = \mathbf{C} = \boldsymbol{\Sigma}_{uu}^{-1}\boldsymbol{\Sigma}_{uw}$. Therefore, given that $\boldsymbol{\Sigma}_{uu}$ is nonsingular, one may test the hypothesis that ordinary least squares is unbiased by testing the hypothesis that $\boldsymbol{\beta}_1 = \mathbf{C}$.

THEOREM 2. *Let the assumptions of Case* (i) *of model* (2.0) *hold and let* $\boldsymbol{\Sigma}_{uu}^{-1}\boldsymbol{\Sigma}_{ue} = \boldsymbol{\beta}_1$. *Let*

$$\hat{\chi}_k^2 = (\tilde{\boldsymbol{\beta}}_1 - \hat{\mathbf{C}})'\tilde{\mathbf{V}}^{-1}(\tilde{\boldsymbol{\beta}}_1 - \hat{\mathbf{C}}),$$

where

$$\hat{\mathbf{C}} = \mathbf{S}_{uu}^{-1}\mathbf{S}_{uw},$$
$$\tilde{\mathbf{V}} = n^{-1}\hat{\mathbf{H}}^{-1}s_v^2 + (n^{-1}s_v^2 + d^{-1}s_r^2)\hat{\mathbf{H}}^{-1}\mathbf{S}_{uu}\hat{\mathbf{H}}^{-1} + d^{-1}(2\hat{\mathbf{H}}^{-1} + \mathbf{S}_{uu}^{-1})s_r^2,$$
$$s_v^2 = (n - k)^{-1}(\mathbf{Y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_1)'(\mathbf{Y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_1),$$
$$s_r^2 = (d - k)^{-1}(S_{ww} - \mathbf{S}_{wu}\mathbf{S}_{uu}^{-1}\mathbf{S}_{uw})$$

and $\hat{\mathbf{H}}$ *is defined following* (3.1). *Then* $\tilde{\chi}_k^2$ *converges in distribution to a chi-square random variable with k degrees of freedom.*

PROOF. By arguments completely analogous to those of Theorem 1, the vector $n^{\frac{1}{2}}[(\tilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1)', (\hat{\mathbf{C}} - \mathbf{C})']'$ converges in distribution to a multivariate normal random variable. The limiting covariance matrix for $n^{\frac{1}{2}}(\tilde{\boldsymbol{\beta}}_1 - \hat{\mathbf{C}})$ under the null ($\boldsymbol{\Sigma}_{uv} = \mathbf{0}$) is

$$\overline{\mathbf{M}}_{xx}^{-1}\sigma_v^2 + \overline{\mathbf{M}}_{xx}^{-1}\boldsymbol{\Sigma}_{uu}\overline{\mathbf{M}}_{xx}^{-1}(\sigma_v^2 + \eta\sigma_r^2) + \eta(2\overline{\mathbf{M}}_{xx}^{-1} + \boldsymbol{\Sigma}_{uu}^{-1})\sigma_r^2,$$

and $n\tilde{\mathbf{V}}$ is a consistent estimator of this matrix. □

In obtaining our results we assumed $(e_t, w_t, \mathbf{u}_t)$ to be multivariate normal. The moment properties of the normal distribution enabled us to obtain explicit expressions for the covariance matrix of the limiting distribution. It is clear from the proof of Theorem 1 that the estimator $\hat{\boldsymbol{\beta}}_1$ will be normally distributed in the limit for $(e_t, w_t, \mathbf{u}_t)$ with finite fourth moments and estimators $\mathbf{S}$ that converge in distribution to normality.

4. Estimation for Case (ii) and (iia). Case (ii) is very similar to Case (i). However, the results follow immediately from those of Case (i) only if $d = \infty$.

THEOREM 3. *Let the assumptions of Case* (ii) *of model* (2.0) *hold. Let the estimator* $\hat{\boldsymbol{\beta}}_1$ *be defined by*

$$\hat{\boldsymbol{\beta}}_1 = (\hat{\mathbf{H}} + n^{-1}\alpha\mathbf{S}_{uu})^{-1}\hat{\mathbf{M}}_{XY}, \tag{4.1}$$

where $\alpha > 0$ *is a fixed number and* $\hat{\mathbf{H}}$ *is given in* (3.1). *Then* $n^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1)$ *converges in distribution to a normal random variable with zero mean and covariance matrix*

$$\overline{\mathbf{M}}_{xx}^{-1}\sigma_v^2 + \overline{\mathbf{M}}_{xx}^{-1}[\boldsymbol{\Sigma}_{uu}(\sigma_v^2 + \eta\boldsymbol{\beta}_1'\boldsymbol{\Sigma}_{uu}\boldsymbol{\beta}_1) + (1 + \eta)\boldsymbol{\Sigma}_{uv}\boldsymbol{\Sigma}_{vu}]\overline{\mathbf{M}}_{xx}^{-1}.$$

PROOF. We have

$$\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1 = \mathbf{M}_{xx}^{-1}(\mathbf{f} - \mathbf{a}\boldsymbol{\beta}_1) + \mathcal{O}_p(n^{-1}),$$

where

$$\mathbf{a} = n^{-1}(\mathbf{x}'\mathbf{u} + \mathbf{u}'\mathbf{x} + \mathbf{u}'\mathbf{u}) - (1 - \alpha n^{-1})\mathbf{S}_{uu}$$

$$\mathbf{f} = n^{-1}(\mathbf{u}'\mathbf{y} + \mathbf{x}'\mathbf{e} + \mathbf{u}'\mathbf{e}),$$

$$\mathbf{f} - \mathbf{a}\boldsymbol{\beta}_1 = n^{-1}(\mathbf{x}'\mathbf{v} + \mathbf{u}'\mathbf{v}) + (1 - \alpha n^{-1})\mathbf{S}_{uu}\boldsymbol{\beta}_1.$$

Evaluating $E\{(\mathbf{f} - \mathbf{a}\boldsymbol{\beta}_1)(\mathbf{f} - \mathbf{a}\boldsymbol{\beta}_1)'\}$ we obtain the variance result. Considering the random variables

$$\Sigma_{i=1}^k \lambda_i \{x_{it}v_t + (u_{it}v_t - \sigma_{u_iv})\}$$

and

$$d^{\frac{1}{2}}\Sigma_{i=1}^k \lambda_i \Sigma_{j=1}^k (s_{ij} - \sigma_{ij})\beta_j,$$

where s_{ij} is the ijth element of $\mathbf{S}_{uu}$ and β_j is the jth element of $\boldsymbol{\beta}_1$, normality follows as in Theorem 1. □

As in Case (i), $\alpha = k + 4 + 2\eta$ gives a smaller mean square error for $\hat{\boldsymbol{\beta}}_1$ than does any smaller α.

In Case (iia) it is known that $\boldsymbol{\Sigma}_{uu} = \text{diag}(\sigma_{11}, \sigma_{22}, \cdots, \sigma_{kk})$ and it is natural to use the matrix $\mathbf{D}_{uu} = \text{diag}(s_{11}, s_{22}, \cdots, s_{kk})$, where s_{ii} is based upon d_i degrees of freedom, in the estimation of $\boldsymbol{\beta}_1$. We consider the estimator

$$(4.2) \qquad \tilde{\boldsymbol{\beta}}_1 = (\hat{\mathbf{H}} + n^{-1}\alpha\mathbf{D}_{uu})^{-1}\hat{\mathbf{M}}_{XY},$$

where $\hat{\mathbf{H}}$ is as defined in (3.1) with $\mathbf{S} = \text{diag}(S_{ww}, s_{11}, s_{22}, \cdots, s_{kk})$ when S_{ww} is known and $\mathbf{S} = \text{diag}(0, s_{11}, s_{22}, \cdots, s_{kk})$ when S_{ww} is unknown.

THEOREM 4. *Let the assumptions of Case* (iia) *of model* (2.0) *hold. Then* $n^{\frac{1}{2}}(\tilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1)$ *converges in distribution to a normal random variable with mean zero and variance*

$$\overline{\mathbf{M}}_{xx}^{-1}\sigma_v^2 + \overline{\mathbf{M}}_{xx}^{-1}(\boldsymbol{\Sigma}_{uu}\sigma_v^2 + \boldsymbol{\Sigma}_{uv}\boldsymbol{\Sigma}_{vu} + 2\mathbf{R})\overline{\mathbf{M}}_{xx}^{-1},$$

where $\mathbf{R} = \text{diag}(\eta_1\beta_1^2\sigma_{11}^2, \eta_2\beta_2^2\sigma_{22}^2, \cdots, \eta_k\beta_k^2\sigma_{kk}^2)$, *and* $\boldsymbol{\beta}_1' = (\beta_1, \beta_2, \cdots, \beta_k)$.

PROOF. The proof parallels that for Case (ii). We have

$$\tilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1 = \mathbf{M}_{xx}^{-1}[n^{-1}(\mathbf{x}'\mathbf{v} + \mathbf{u}'\mathbf{v}) + \mathbf{D}_{uu}\boldsymbol{\beta}_1] + \mathcal{O}_p(n^{-1})$$

and the variance result follows from

$$E\{(\mathbf{D}_{uu}\boldsymbol{\beta}_1 + \boldsymbol{\Sigma}_{uv})(\mathbf{D}_{uu}\boldsymbol{\beta}_1 + \boldsymbol{\Sigma}_{uv})'\} = 2n^{-1}\mathbf{R}. \qquad \square$$

Comparison of the variance formulas for Cases (i), (ii) and (iia) makes it clear that the use of the knowledge that some of the measurement error covariances are zero reduces the variance of the estimator of $\boldsymbol{\beta}_1$. The variance of the estimator of $\boldsymbol{\beta}_1$ is greater for Case (i) than for Case (ii) when $\boldsymbol{\Sigma}_{ue} = \mathbf{0}$ because

$$\boldsymbol{\Sigma}_{uu}\sigma_r^2 - \boldsymbol{\Sigma}_{uu}(\boldsymbol{\beta}_1'\boldsymbol{\Sigma}_{uu}\boldsymbol{\beta}_1)$$

is then a positive semidefinite matrix. The variance for Case (ii) is greater than that of Case (iia) because, for $\boldsymbol{\Sigma}_{uu}$ diagonal and $\eta = \eta_1 = \eta_2 = \cdots = \eta_k$,

$$\eta\boldsymbol{\Sigma}_{uu}(\boldsymbol{\beta}_1'\boldsymbol{\Sigma}_{uu}\boldsymbol{\beta}_1) - \mathbf{R}$$

is a positive semidefinite matrix.

The reader will note that, with the exception of Theorem 2, there is nothing in the proofs of the theorems that precludes the specification that some $\sigma_{ii} = 0$. Naturally if $\sigma_{ii} = 0$ the ith row and column of $\mathbf{S}_{uu}$ will be zero. The results of Theorem 2 can be extended to singular $\boldsymbol{\Sigma}_{uu}$ matrices by suitable reparameterization.

5. Estimation for Case (iii). In Case (iii) the expected value of the estimator of $\boldsymbol{\Sigma}$ is permitted to differ from $\boldsymbol{\Sigma}$ by a multiple. We assume the estimator is distributed as a multiple of a Wishart matrix with d degrees of freedom, independent of $\mathbf{X}$ and $\mathbf{Y}$. Denoting the estimator by $\tilde{\mathbf{S}}$ we have

$$E\{\tilde{\mathbf{S}}\} = \lambda^{-1}\boldsymbol{\Sigma},$$

where $\tilde{\mathbf{S}}$ is partitioned as

$$\tilde{\mathbf{S}} = \begin{pmatrix} \tilde{S}_{ee} & \tilde{\mathbf{S}}_{eu} \\ \tilde{\mathbf{S}}_{ue} & \tilde{\mathbf{S}}_{uu} \end{pmatrix},$$

and we define $\mathbf{S} = \lambda\tilde{\mathbf{S}}$. Anderson (1951) demonstrated that the maximum likelihood estimator of $\boldsymbol{\beta}_1$, denoted by $\hat{\boldsymbol{\beta}}_{L1}$, is given by

$$(\hat{\mathbf{M}}_{XX} - \hat{\lambda}\tilde{\mathbf{S}}_{uu})\hat{\boldsymbol{\beta}}_{L1} = \hat{\mathbf{M}}_{XY} - \hat{\lambda}\tilde{\mathbf{S}}_{ue}, \tag{5.1}$$

where $\hat{\lambda}$ is the smallest root of

$$|\hat{\mathbf{M}} - \lambda\tilde{\mathbf{S}}| = 0 \tag{5.2}$$

and $\hat{\mathbf{M}}$ was defined in Section 3. If we define the vector $\hat{\boldsymbol{\beta}}_L$ by $\hat{\boldsymbol{\beta}}_L = [1, -\hat{\boldsymbol{\beta}}'_{L1}]'$, then

$$\hat{\lambda} = (\hat{\boldsymbol{\beta}}'_L\tilde{\mathbf{S}}\hat{\boldsymbol{\beta}}_L)^{-1}\hat{\boldsymbol{\beta}}'_L\hat{\mathbf{M}}\hat{\boldsymbol{\beta}}_L$$

and

$$(\hat{\mathbf{M}} - \hat{\lambda}\tilde{\mathbf{S}})\hat{\boldsymbol{\beta}}_L = 0. \tag{5.3}$$

Anderson (1948), following Hsu (1941), gave the limiting distribution of the roots of (5.2) for $\tilde{\mathbf{S}} = \boldsymbol{\Sigma}$. We demonstrate for our model that the distribution of $\hat{\lambda}$ can be approximated by a multiple of an F random variable.

THEOREM 5. *Let the assumptions of Case* (iii) *of model* (2.0) *hold. Then*

$$(n - k)^{-1}n\hat{\lambda} = \lambda F + \mathcal{O}_p(n^{-1}),$$

where F is a random variable distributed as Snedecor's F with $n - k$ and d degrees of freedom.

PROOF. Because $\hat{\lambda}$ is a continuous function of the elements of $\hat{\mathbf{M}}$ and $\tilde{\mathbf{S}}$ in an open sphere containing the true values and because, by our assumptions,

$$p \lim \hat{\mathbf{M}} = \overline{\mathbf{M}}_{zz} + \boldsymbol{\Sigma},$$
$$p \lim \tilde{\mathbf{S}} = \lambda^{-1}\boldsymbol{\Sigma},$$

it follows that $p \lim \hat{\lambda} = \lambda$. Likewise

$$p \lim \hat{\boldsymbol{\beta}}_{L1} = p \lim\left\{(\hat{\mathbf{M}}_{XX} - \hat{\lambda}\tilde{\mathbf{S}}_{uu})^{-1}(\hat{\mathbf{M}}_{XY} - \hat{\lambda}\tilde{\mathbf{S}}_{ue})\right\} = \boldsymbol{\beta}_1.$$

Defining

$$\begin{aligned}
\Delta\mathbf{M} &= \hat{\mathbf{M}} - E\{\hat{\mathbf{M}}\} = \hat{\mathbf{M}} - \mathbf{M}, \\
\Delta\tilde{\mathbf{S}} &= \tilde{\mathbf{S}} - \lambda^{-1}\boldsymbol{\Sigma}, \\
\Delta\lambda &= \hat{\lambda} - \lambda, \\
\Delta\boldsymbol{\beta} &= \hat{\boldsymbol{\beta}}_L - \boldsymbol{\beta} = \left[0, (\boldsymbol{\beta}_1 - \hat{\boldsymbol{\beta}}_{L1})'\right]' = \left[0, -\Delta\boldsymbol{\beta}'_1\right]',
\end{aligned}$$

we write (5.3) as

$$\left[\mathbf{M}_{zz} + \{(\Delta\mathbf{M}) - \lambda(\Delta\tilde{\mathbf{S}})\} - (\Delta\lambda)\lambda^{-1}\boldsymbol{\Sigma} - (\Delta\lambda)(\Delta\tilde{\mathbf{S}})\right]\hat{\boldsymbol{\beta}}_L = \mathbf{0}, \tag{5.4}$$

where we have used $\mathbf{M} = \mathbf{M}_{zz} + \boldsymbol{\Sigma}$. If we premultiply (5.4) by $\boldsymbol{\beta}'$ and note that $\boldsymbol{\beta}'\mathbf{M}_{zz} = \mathbf{0}$ we have

$$(5.5) \qquad \boldsymbol{\beta}'[(\Delta\mathbf{M}) - \lambda(\Delta\tilde{\mathbf{S}})]\hat{\boldsymbol{\beta}}_L - (\Delta\lambda)\lambda^{-1}\boldsymbol{\beta}'\boldsymbol{\Sigma}\hat{\boldsymbol{\beta}}_L - (\Delta\lambda)\boldsymbol{\beta}'(\Delta\tilde{\mathbf{S}})\hat{\boldsymbol{\beta}}_L = 0.$$

As we have proven that $\Delta\boldsymbol{\beta} = o_p(1)$,

$$\Delta\lambda = (\boldsymbol{\beta}'\tilde{\mathbf{S}}\boldsymbol{\beta})^{-1}\boldsymbol{\beta}'[(\Delta\mathbf{M}) - \lambda(\Delta\tilde{\mathbf{S}})]\boldsymbol{\beta} + o_p(n^{-\frac{1}{2}}).$$

Using (5.1) and the definition of $\Delta\mathbf{M}$, we obtain

$$(5.6) \quad \hat{\boldsymbol{\beta}}_{L1} - \boldsymbol{\beta}_1 = \mathbf{M}_{xx}^{-1}[n^{-1}(\mathbf{x}'\mathbf{v} + \mathbf{u}'\mathbf{v}) - \lambda\tilde{\mathbf{S}}_{uv} - \lambda^{-1}(\Delta\lambda)\boldsymbol{\Sigma}_{uv}] + \mathcal{O}_p(n^{-1}),$$

where $\mathbf{v} = \mathbf{e} - \mathbf{u}\boldsymbol{\beta}_1 = \boldsymbol{\varepsilon}\boldsymbol{\beta}$. From (5.5)

$$\hat{\lambda} = \frac{\boldsymbol{\beta}'[(\Delta\mathbf{M}) + \boldsymbol{\Sigma}]\boldsymbol{\beta}}{\boldsymbol{\beta}'\tilde{\mathbf{S}}\boldsymbol{\beta}} + \frac{\boldsymbol{\beta}'[(\Delta\mathbf{M}) - \lambda(\Delta\mathbf{S})](\Delta\boldsymbol{\beta})}{\boldsymbol{\beta}'\tilde{\mathbf{S}}\boldsymbol{\beta}} - \frac{\boldsymbol{\beta}'[(\Delta\mathbf{M}) - \lambda(\Delta\mathbf{S})]\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{S}(\Delta\boldsymbol{\beta})}{(\boldsymbol{\beta}'\tilde{\mathbf{S}}\boldsymbol{\beta})^2}$$

$$+ \mathcal{O}_p(n^{-\frac{3}{2}}) = (n\boldsymbol{\beta}'\tilde{\mathbf{S}}\boldsymbol{\beta})^{-1}\mathbf{v}'(\mathbf{I} - n^{-1}\mathbf{x}\mathbf{M}_{xx}^{-1}\mathbf{x}')\mathbf{v} + \mathcal{O}_p(n^{-1}).$$

By assumption, $\boldsymbol{\beta}'\tilde{\mathbf{S}}\boldsymbol{\beta}$ is independent of $\mathbf{v} = \boldsymbol{\varepsilon}\boldsymbol{\beta}$. Furthermore $\boldsymbol{\beta}'\tilde{\mathbf{S}}\boldsymbol{\beta}$ is distributed as a multiple of a chi-square random variable with d degrees of freedom, where the multiple is $d^{-1}\lambda^{-1}\sigma_v^2$. □

Because of the analogy to linear regression theory, we chose to include the term $n^{-2}\mathbf{v}'\mathbf{x}\mathbf{M}_{xx}^{-1}\mathbf{x}'\mathbf{v}$ in the numerator of the F-statistic although this term is $\mathcal{O}_p(n^{-1})$. In practice, when λ is known and the value of $\lambda^{-1}\hat{\lambda}$ is large (or small) relative to the critical value of the central F distribution, the validity of the model is called into question.

THEOREM 6. *Let the assumptions of Case* (iii) *of model* (2.0) *hold. Then* $n^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}_{L1} - \boldsymbol{\beta}_1)$ *converges in distribution to a normal random variable with mean zero and covariance matrix*

$$\overline{\mathbf{M}}_{xx}^{-1}\sigma_v^2 + (1 + \eta)\overline{\mathbf{M}}_{xx}^{-1}(\boldsymbol{\Sigma}_{uu}\sigma_v^2 - \boldsymbol{\Sigma}_{uv}\boldsymbol{\Sigma}_{vu})\overline{\mathbf{M}}_{xx}^{-1}.$$

PROOF. By using (5.5) and

$$\Delta\lambda = \lambda\sigma_v^{-2}(\hat{\sigma}_v^2 - s_v^2) + \mathcal{O}_p(n^{-1}),$$

where $\hat{\sigma}_v^2 = n^{-1}\mathbf{v}'\mathbf{v}$ and $s_v^2 = \boldsymbol{\beta}'\tilde{\mathbf{S}}\boldsymbol{\beta}$, we have

$$\hat{\boldsymbol{\beta}}_{L1} - \boldsymbol{\beta}_1 = \mathbf{M}_{xx}^{-1}[n^{-1}(\mathbf{x}'\mathbf{v} + \mathbf{u}'\mathbf{v}) - \tilde{\mathbf{S}}_{uv} - \sigma_v^{-2}(\hat{\sigma}_v^2 - s_v^2)\boldsymbol{\Sigma}_{uv}] + \mathcal{O}_p(n^{-1}).$$

Considering the random variables

$$\Sigma_{i=1}^k C_i\{x_{it}v_t + u_{it}v_t - \sigma_v^{-2}v_t^2\sigma_{u_iv}\}, \qquad t = 1, 2, \cdots, n,$$

and

$$\Sigma_{i=1}^k C_i(s_{u_iv} - \sigma_v^{-2}s_v^2\sigma_{u_iv}),$$

where the C_i are arbitrary real numbers, asymptotic normality follows by the argument of Theorem 1. □

The root $\hat{\lambda}$ is bounded by a multiple of an F random variable and, hence, has finite moments for sufficiently large d. However, $\hat{\beta}_{L1}$ does not necessarily possess finite moments. Therefore we suggest that the maximum likelihood estimator be replaced by

$$\tilde{\boldsymbol{\beta}}_{L1} = \left[\hat{\mathbf{M}}_{XX} - \hat{\lambda}(1 - n^{-1}\alpha)\tilde{\mathbf{S}}_{uu}\right]^{-1}\left[\hat{\mathbf{M}}_{XY} - \hat{\lambda}(1 - n^{-1}\alpha)\tilde{\mathbf{S}}_{ue}\right],$$

where $\alpha > 0$ is fixed. Under the assumption of bounded x's it has been demonstrated that

$$E\{\tilde{\boldsymbol{\beta}}_{L1} - \boldsymbol{\beta}_1\} = -\Big[n^{-1}(1-\alpha)\mathbf{I} + (n^{-1} + d^{-1})$$
$$\times \mathbf{M}_{xx}^{-1}\{\boldsymbol{\Sigma}_{uu} - (\sigma_v^2)^{-1}\boldsymbol{\Sigma}_{uv}\boldsymbol{\Sigma}_{vu}\}\Big]\mathbf{M}_{xx}^{-1} + \mathcal{O}(n^{-2}).$$

Also it is possible to prove that, through terms of $\mathcal{O}(n^{-2})$, the mean square error of $\tilde{\boldsymbol{\beta}}_{L1}$ is uniformly smaller for $\alpha = 4$ than for any smaller α.

Case (iiia) differs from Case (iii) in that

$$\boldsymbol{\Sigma} = \operatorname{diag}(\sigma_{00}, \sigma_{11}, \cdots, \sigma_{kk})$$

is estimated by

$$\mathbf{D} = \operatorname{diag}(s_{00}, s_{11}, \cdots, s_{kk}),$$

where the s_{ii} are assumed to be independently distributed as multiples of chi-square random variables with d_i, $i = 0, 1, \cdots, k$, degrees of freedom. In our presentation we assume $E\{\mathbf{D}\} = \boldsymbol{\Sigma}$, but it is easy to modify the results for the case wherein $E\{\mathbf{D}\} = \lambda^{-1}\boldsymbol{\Sigma}$.

We consider the estimator

$$\tilde{\boldsymbol{\beta}}_{L1} = \left[\hat{\mathbf{M}}_{XX} - (\hat{\lambda} - n^{-1}\alpha)\mathbf{D}_{uu}\right]^{-1}\hat{\mathbf{M}}_{XY},$$

where $\alpha > 0$ is fixed and $\hat{\lambda}$ is the smallest root of $|\hat{M} - \hat{\lambda}D| = 0$.

THEOREM 7. *Let the assumptions of Case* (iiia) *of model* (2.0) *hold. Let* $E\{\mathbf{D}\} = \boldsymbol{\Sigma}$. *Then* $n^{\frac{1}{2}}(\hat{\lambda} - \lambda)$ *converges in distribution to a normal random variable with mean zero and variance*

$$2 + 2\sigma_v^{-4}\Sigma_{i=0}^{k}\eta_i\beta_i^4\sigma_{ii}^2,$$

where $\boldsymbol{\beta}' = (1, -\boldsymbol{\beta}_1') = (\beta_0, -\beta_1, -\beta_2, \cdots, -\beta_k)$.

PROOF. As in Theorem 5, $\hat{\lambda}$ is bounded by a multiple of an F random variable. Using (5.4)

$$\Delta\lambda = \frac{\boldsymbol{\beta}'[(\Delta\mathbf{M}) - (\Delta\mathbf{D})]\boldsymbol{\beta}}{\sigma_v^2} + \mathcal{O}_p(n^{-1}),$$

where

$$\Delta\mathbf{D} = \mathbf{D} - \boldsymbol{\Sigma},$$
$$\boldsymbol{\beta}'(\Delta\mathbf{M})\boldsymbol{\beta} = n^{-1}\mathbf{v}'\mathbf{v} - \sigma_v^2,$$
$$\boldsymbol{\beta}'(\Delta\mathbf{D})\boldsymbol{\beta} = \Sigma_{i=0}^{k}\beta_i^2 s_{ii} - \sigma_v^2.$$

Hence

$$\Delta\lambda = \sigma_v^2\left(n^{-1}\mathbf{v}'\mathbf{v} - \Sigma_{i=0}^k \beta_i^2 s_{ii}\right) + \mathcal{O}_p(n^{-1})$$

and the result follows by the arguments of Theorem 1. □

By analogy to Theorem 5 one might choose to approximate the distribution of $(n-k)^{-1}n\hat{\lambda}$ with the distribution of Snedecor's F with $n-k$ and ν degrees of freedom, where

$$\nu = \left[\Sigma_{i=0}^k d_i^{-1}\beta_i^4\sigma_{ii}^2\right]^{-1}\sigma_v^4.$$

THEOREM 8. *Let the assumptions of Case* (iiia) *of model* (2.0) *hold. Then* $n^{\frac{1}{2}}(\tilde{\boldsymbol{\beta}}_{L1} - \boldsymbol{\beta}_1)$ *converges in distribution to a normal random variable with mean zero and covariance matrix*

$$\overline{\mathbf{M}}_{xx}^{-1}\sigma_v^2 + \overline{\mathbf{M}}_{xx}^{-1}(\boldsymbol{\Sigma}_{uu}\sigma_v^2 - \boldsymbol{\Sigma}_{uv}\boldsymbol{\Sigma}_{vu})\overline{\mathbf{M}}_{xx}^{-1}$$

$$+2\overline{\mathbf{M}}_{xx}^{-1}\left\{\mathbf{R} + \boldsymbol{\Sigma}_{uv}\boldsymbol{\Sigma}_{vu}\sigma_v^{-4}\left(\Sigma_{i=0}^k\eta_i\beta_i^4\sigma_{ii}^2\right) + \sigma_v^{-2}(\mathbf{R}\boldsymbol{\beta}_1\boldsymbol{\Sigma}_{vu} + \boldsymbol{\Sigma}_{uv}\boldsymbol{\beta}_1'\mathbf{R})\right\}\overline{\mathbf{M}}_{xx}^{-1},$$

where

$$\mathbf{R} = \operatorname{diag}\left(\eta_1\beta_1^2\sigma_{11}^2, \eta_2\beta_2^2\sigma_{22}^2, \cdots, \eta_k\beta_k^2\sigma_{kk}^2\right).$$

PROOF. We have

$$\begin{aligned}\tilde{\boldsymbol{\beta}}_{L1} - \boldsymbol{\beta}_1 &= \mathbf{M}_{xx}^{-1}\left[n^{-1}(\mathbf{x}'\mathbf{v} + \mathbf{u}'\mathbf{v}) + \mathbf{D}_{uu}\boldsymbol{\beta}_1 + (\Delta\lambda)\boldsymbol{\Sigma}_{uu}\boldsymbol{\beta}_1\right] + \mathcal{O}_p(n^{-1})\\ &= \mathbf{M}_{xx}^{-1}\left[n^{-1}(\mathbf{x}'\mathbf{v} + \mathbf{u}'\mathbf{v}) + \mathbf{D}_{uu}\boldsymbol{\beta}_1 - \sigma_v^{-2}(\hat{\sigma}_v^2 - \Sigma_{i=0}^k\beta_i^2 s_{ii})\boldsymbol{\Sigma}_{uv}\right] + \mathcal{O}_p(n^{-1}).\end{aligned}$$

Considering the random variables

$$\Sigma_{i=1}^k C_i\left\{x_{it}v_t + u_{it}v_t - \sigma_v^{-2}v_t^2\sigma_{u_iv}\right\}, \qquad t = 1, 2, \cdots, n,$$

and

$$\Sigma_{i=1}^k C_i\left\{s_{ii}\beta_i + \sigma_v^{-2}\Sigma_{j=0}^k\beta_j^2 s_{jj}\sigma_{u_iv}\right\},$$

where the C_i are arbitrary real constants, the result follows by the arguments of Theorem 1. □

APPENDIX

THEOREM A. *Let the assumptions of Case* (i) *of model* (2.0) *hold and let the sequence* $\{|\mathbf{x}_{t.}|\}$ *be uniformly bounded. Then*

$$E\left\{\tilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1\right\} = -n^{-1}\mathbf{M}_{xx}^{-1}\left\{(k+1-\alpha)\mathbf{I} + [1+\eta]\left[\operatorname{tr}(\boldsymbol{\Sigma}_{uu}\mathbf{M}_{xx}^{-1})\mathbf{I} + \boldsymbol{\Sigma}_{uu}\mathbf{M}_{xx}^{-1}\right]\right\}\boldsymbol{\Sigma}_{uv} + \mathcal{O}(n^{-2}).$$

Furthermore, through terms of order n^{-2}, *the mean square error of* $\tilde{\boldsymbol{\beta}}_1$ *is smaller for* $\alpha = k + 4 + 2\eta$ *than for any smaller* α.

PROOF. The ith element of $\hat{\boldsymbol{\beta}}_1$ is given by

$$\hat{\beta}_i = \Sigma_{j=1}^k \hat{h}^{ij}\hat{N}_j = \Sigma_{j=1}^k |\hat{\mathbf{H}}|^{-1}\operatorname{cof}(\hat{h}_{ij})\hat{N}_j$$

where cof $(\hat{h}_{ij})$ is the signed cofactor of $\hat{h}_{ij}$, $\hat{h}_{ij}$ is the ijth element of $\hat{\mathbf{H}}$, $\hat{h}^{ij}$ is the ijth element of $\hat{\mathbf{H}}^{-1}$, and $\hat{N}_j$ is the jth element of $\hat{\mathbf{N}}$. Let $\hat{\boldsymbol{\theta}}$ denote the column vector obtained by listing the columns of $\hat{\mathbf{M}}_{xx}$, $\hat{\mathbf{M}}_{xy}$, $\mathbf{S}_{uu}$, $\mathbf{S}_{ue}$ in a single column and let $L(\hat{\boldsymbol{\theta}}) = |\mathbf{S}_{uu}|^{-r}|\Sigma_{j=1}^{k}\text{cof}(\hat{h}_{ij})\hat{N}_j|^r$ where r is a positive integer. Because

$$|\hat{\mathbf{H}}| \geqslant |n^{-1}\mathbf{S}_{uu}|$$

we have

$$|\hat{\beta}_i|^r \leqslant n^{rk}L(\hat{\boldsymbol{\theta}}).$$

James (1954) has shown that $E\{|\mathbf{S}_{uu}|^{-4rk}\}$ is bounded for all d greater than some number depending on r. By the normality of $\boldsymbol{\varepsilon}$ and because $|\mathbf{x}_{t.}|$ is bounded,

$$E\{|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}|^{4rk}\} = \mathcal{O}(n^{-2rk})$$

for $r = 1, 2, \cdots$, where $\boldsymbol{\theta} = E\{\hat{\boldsymbol{\theta}}\}$. It follows that, for $r = 1, 2, \cdots$,

$$E\{|L(\hat{\boldsymbol{\theta}})|^{4k}\} = \mathcal{O}(1).$$

We may choose an open set A containing $\boldsymbol{\theta}$ such that there exists an n_1 for which $\hat{\gamma} > 1 + n^{-1}$ if $\hat{\boldsymbol{\theta}} \in A$ and $n > n_1$. For $n > n_1$ and $\hat{\boldsymbol{\theta}} \in A$ the elements of $\tilde{\boldsymbol{\beta}}_1$ are continuous functions of $\hat{\boldsymbol{\theta}}$ with continuous third derivatives. Therefore the conditions of Theorems 5.4.3 and 5.4.4 of Fuller (1976) are satisfied and a truncated Taylor's series may be used to obtain the moments of $\tilde{\boldsymbol{\beta}}_1$ through terms of order n^{-2}. Carrying out the Taylor's expansion we have

$$\tilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1 = \mathbf{M}_{xx}^{-1}(\mathbf{b} - \mathbf{a}\boldsymbol{\beta}_1) - \mathbf{M}_{xx}^{-1}\mathbf{a}\mathbf{M}_{xx}^{-1}(\mathbf{b} - \mathbf{a}\boldsymbol{\beta}_1) + \mathcal{O}_p(n^{-\frac{3}{2}}),$$

where

$$\mathbf{a} = n^{-1}(\mathbf{x}'\mathbf{u} + \mathbf{u}'\mathbf{x} + \mathbf{u}'\mathbf{u}) - (1 - \alpha n^{-1})\mathbf{S}_{uu},$$
$$\mathbf{b} = n^{-1}(\mathbf{u}'\mathbf{y} + \mathbf{x}'\mathbf{e} + \mathbf{u}'\mathbf{e}) - (1 - \alpha n^{-1})\mathbf{S}_{uw}.$$

Using

$$n^{-1}E\{\mathbf{x}'\mathbf{u}\mathbf{M}_{xx}^{-1}\mathbf{x}'\mathbf{v}\} = \boldsymbol{\Sigma}_{uv},$$
$$n^{-1}E\{\mathbf{u}'\mathbf{x}\mathbf{M}_{xx}^{-1}\mathbf{x}'\mathbf{v}\} = k\boldsymbol{\Sigma}_{uv},$$
$$E\{(n^{-1}\mathbf{u}'\mathbf{u} - \mathbf{S}_{uu})\mathbf{M}_{xx}^{-1}(n^{-1}\mathbf{u}'\mathbf{v} - \mathbf{S}_{ur})\} = (n^{-1} + d^{-1})[\text{tr}(\boldsymbol{\Sigma}_{uu}\mathbf{M}_{xx}^{-1})\mathbf{I} + \boldsymbol{\Sigma}_{uu}\mathbf{M}_{xx}^{-1}]\boldsymbol{\Sigma}_{uv},$$

we obtain the bias result. Similarly,

$$\begin{aligned}(\tilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1)(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_1)' = {} & (\hat{\mathbf{M}}_{XX} - \mathbf{S}_{uu})^{-1}\mathbf{g}\mathbf{g}'(\hat{\mathbf{M}}_{XX} - \mathbf{S}_{uu})^{-1} \\ & - n^{-1}\alpha\mathbf{M}_{xx}^{-1}\boldsymbol{\Sigma}_{uu}\mathbf{M}_{xx}^{-1}\mathbf{g}\mathbf{g}'\mathbf{M}_{xx}^{-1} - n^{-1}\alpha\mathbf{M}_{xx}^{-1}\mathbf{g}\mathbf{g}'\mathbf{M}_{xx}^{-1}\boldsymbol{\Sigma}_{uu}\mathbf{M}_{xx}^{-1} \\ & + n^{-1}\alpha\mathbf{M}_{xx}^{-1}(\mathbf{g}\mathbf{S}_{ru} + \mathbf{S}_{ur}\mathbf{g}')\mathbf{M}_{xx}^{-1} \\ & - n^{-1}\alpha\mathbf{M}_{xx}^{-1}\mathbf{a}\mathbf{M}_{xx}^{-1}(\mathbf{g}\boldsymbol{\Sigma}_{vu} + \boldsymbol{\Sigma}_{uv}\mathbf{g}')\mathbf{M}_{xx}^{-1} \\ & - n^{-1}\alpha\mathbf{M}_{xx}^{-1}(\mathbf{g}\boldsymbol{\Sigma}_{vu} + \boldsymbol{\Sigma}_{uv}\mathbf{g}')\mathbf{M}_{xx}^{-1}\mathbf{a}\mathbf{M}_{xx}^{-1} \\ & + n^{-2}\alpha^2\mathbf{M}_{xx}^{-1}\boldsymbol{\Sigma}_{uv}\boldsymbol{\Sigma}_{vu}\mathbf{M}_{xx}^{-1} + \mathcal{O}_p(n^{-\frac{5}{2}}),\end{aligned} \tag{A.1}$$

where $\mathbf{g} = n^{-1}(\mathbf{x}'\mathbf{v} + \mathbf{u}'\mathbf{v}) - \mathbf{S}_{ur}$. Now the first term of (A.1) is, except for modification associated with $\hat{\gamma}$, the error in the estimator with $\alpha = 0$. Therefore we need only evaluate the expectation of the remaining six terms of (A.1). Using

$$E\{\mathbf{g}\mathbf{S}_{ru}\} = -d^{-1}(\sigma_r^2\mathbf{\Sigma}_{uu} + \mathbf{\Sigma}_{uv}\mathbf{\Sigma}_{vu})$$

$$E\{\mathbf{c}\mathbf{M}_{xx}^{-1}\mathbf{g}\mathbf{\Sigma}_{vu}\} = n^{-1}(k+1)\mathbf{\Sigma}_{uv}\mathbf{\Sigma}_{vu} + (n^{-1} + d^{-1}) \times [\mathbf{\Sigma}_{uu}\mathbf{M}_{xx}^{-1} + \operatorname{tr}(\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uu})\mathbf{I}]\mathbf{\Sigma}_{uv}\mathbf{\Sigma}_{vu}$$

$$E\{\mathbf{c}\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uv}\mathbf{g}'\} = n^{-1}\mathbf{M}_{xx}\mathbf{\Sigma}_{vu}\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uv} + n^{-1}\mathbf{\Sigma}_{uv}\mathbf{\Sigma}_{vu} + (n^{-1} + d^{-1})[\mathbf{\Sigma}_{uu}(\mathbf{\Sigma}_{vu}\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uv}) + \mathbf{\Sigma}_{uv}\mathbf{\Sigma}_{vu}\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uu}],$$

where $\mathbf{c} = n^{-1}(\mathbf{x}'\mathbf{u} + \mathbf{u}'\mathbf{v} + \mathbf{u}'\mathbf{u}) - \mathbf{S}_{uu}$, we obtain

$$\begin{aligned}
&-\alpha n^{-2}\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uu}\mathbf{M}_{xx}^{-1}[(\sigma_v^2 + \eta\sigma_r^2)\mathbf{\Sigma}_{uu} + (1+\eta)\mathbf{\Sigma}_{uv}\mathbf{\Sigma}_{vu}]\mathbf{M}_{xx}^{-1}\\
&-\alpha n^{-2}\mathbf{M}_{xx}^{-1}[(\sigma_v^2 + \eta\sigma_r^2)\mathbf{\Sigma}_{uu} + (1+\eta)\mathbf{\Sigma}_{uv}\mathbf{\Sigma}_{vu}]\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uu}\mathbf{M}_{xx}^{-1}\\
&-2\alpha n^{-2}(1+\eta)\mathbf{M}_{xx}^{-1}[\mathbf{\Sigma}_{uu}^{-1}\mathbf{M}_{xx}^{-1} + \operatorname{tr}(\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uu})\mathbf{I}]\mathbf{\Sigma}_{uv}\mathbf{\Sigma}_{vu}\mathbf{M}_{xx}^{-1}\\
\text{(A.2)}\qquad &-2\alpha n^{-2}\mathbf{M}_{xx}^{-1}(\mathbf{\Sigma}_{vu}\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uv})\\
&-2\alpha n^{-2}(1+\eta)\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uu}\mathbf{M}_{xx}^{-1}(\mathbf{\Sigma}_{vu}\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uv})\\
&-\alpha n^{-2}(1+\eta)[\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uv}\mathbf{\Sigma}_{vu}\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uu}\mathbf{M}_{xx}^{-1} + \mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uu}\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uv}\mathbf{\Sigma}_{vu}\mathbf{M}_{xx}^{-1}]\\
&-n^{-2}\mathbf{M}_{xx}^{-1}[2\alpha\mathbf{\Sigma}_{uu}(\sigma_v^2 + \eta\sigma_r^2) - \{\alpha^2 - 2\alpha(k+2) - 2\alpha\eta\}\mathbf{\Sigma}_{uv}\mathbf{\Sigma}_{vu}]\mathbf{M}_{xx}^{-1}\\
&+\mathcal{O}(n^{-3})
\end{aligned}$$

as the effect of α on the mean square error of $\tilde{\boldsymbol{\beta}}_1$. Let $\boldsymbol{\xi}$ be an arbitrary real k dimensional column vector with $|\boldsymbol{\xi}| \neq 0$. Then $\boldsymbol{\xi}'\mathbf{L}\boldsymbol{\xi}$, where

$$\mathbf{L} = 2\alpha\mathbf{M}_{xx}^{-1}(\mathbf{\Sigma}_{vu}\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uv}) + 2\alpha\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uu}\mathbf{M}_{xx}^{-1}(\sigma_v^2 + \eta\sigma_r^2) - [\alpha^2 - 2\alpha(k+2) - 2\alpha\eta]\mathbf{M}_{xx}^{-1}\mathbf{\Sigma}_{uv}\mathbf{\Sigma}_{vu}\mathbf{M}_{xx}^{-1},$$

is larger for $\alpha = k + 4 + 2\eta$ than for any smaller α. As all other multipliers of α in (A.2) are negative semidefinite matrices, the conclusion follows. □

REFERENCES

[1] Anderson, T. W. (1948). The asymptotic distributions of the roots of certain determinantal equations. *J. Roy. Statist. Soc. Ser. B* **10** 140–158.

[2] Anderson, T. W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. *Ann. Math. Statist.* **22** 327–351.

[3] Anderson, T. W. (1976). Estimation of linear functional relationships: approximate distributions and connections with simultaneous equations in econometrics. *J. Roy. Statist. Soc. Ser. B* **38** 1–20.

[4] Anderson, T. W. and Rubin, H. (1949). Estimation of the parameter of a single equation in a complete system of stochastic equations. *Ann. Math. Statist.* **20** 46–63.

[5] Anderson, T. W. and Rubin, H. (1950). The asymptotic properties of estimates of the parameters of a single equation in a complete system of stochastic equations. *Ann. Math. Statist.* **21** 570–582.

[6] Battese, G. E., Fuller, W. A. and Hickman, R. D. (1976). Estimation of response variances from interview-reinterview surveys. *J. Indian Soc. Agr. Statist.* **28** 1–14.

[7] Cochran, W. G. (1968). Errors of measurement in statistics. *Technometrics* **10** 637–666.

[8] DeGracie, J. S. and Fuller, W. A. (1972). Estimation of the slope and analysis of covariance when the concomitant variable is measured with error. *J. Amer. Statist. Assoc.* **67** 930–937.

[9] Dorff, M. and Gurland, J. (1961). Estimation of the parameters of a linear functional relation. *J. Roy. Statist. Soc. Ser. B* **23** 160–170.

[10] Fuller, W. A. (1976). *Introduction to Statistical Time Series.* Wiley, New York.

[11] Fuller, W. A. (1978). An affine linear model for the relation between two sets of frequency counts: response to query. *Biometrics* **34** 514–521.

[12] Hsu, P. L. (1941). On the limiting distribution of roots of a determinantal equation. *J. London Math. Soc.* **16** 183–194.

[13] James, A. T. (1954). Normal multivariate analysis and the orthogonal group. *Ann. Math. Statist.* **25** 40–75.

[14] Johnston, J. (1972). *Econometric Methods.* McGraw Hill, New York.

[15] Kendall, M. G. and Stuart, A. (1961). *The Advanced Theory of Statistics.* Ch. 27. Griffin, London.

[16] Koopmans, T. C. (1937). *Linear Regression Analysis in Economic Time Series.* Haarlem.

[17] Lawley, D. N. (1953). A modified method of estimation in factor analysis and some large sample results. In *Factor Analysis.* Selected Publications for the Uppsala Univ. Institute of Statistics, Vol. 9.

[18] Lord, F. M. (1960). Large-sample covariance analysis when the control variable is fallible. *J. Amer. Statist. Assoc.* **55** 307–321.

[19] Madansky, W. (1959). The fitting of straight lines when both variables are subject to error. *J. Amer. Statist. Assoc.* **54** 173–205.

[20] Malinvaud, E. (1970). *Statistical Methods of Econometrics.* North-Holland, Amsterdam.

[21] Moran, P. A. P. (1971). Estimating structural and functional relationships. *J. Multivariate Anal.* **1** 232–235.

[22] Rao, C. R. (1965). *Linear Statistical Inference and its Applications.* Wiley, New York.

[23] Schneeweiss, H. (1976). Consistent estimation of a regression with errors in the variables. *Metrika* **23** 101–115.

[24] Villegas, C. (1961). Maximum likelihood estimation of a linear functional relationship. *Ann. Math. Statist.* **32** 1048–1062.

[25] Villegas, C. (1966). On the asymptotic efficiency of least squares estimators. *Ann. Math. Statist.* **37** 1676–1683.

[26] Warren, R. D., Keller, J. P., and Fuller, W. A. (1974). An errors-in-variables analysis of managerial role performance. *J. Amer. Statist. Assoc.* **69** 886–893.

Department of Statistics
Iowa State University
Snedecor Hall
Ames, Iowa 50011