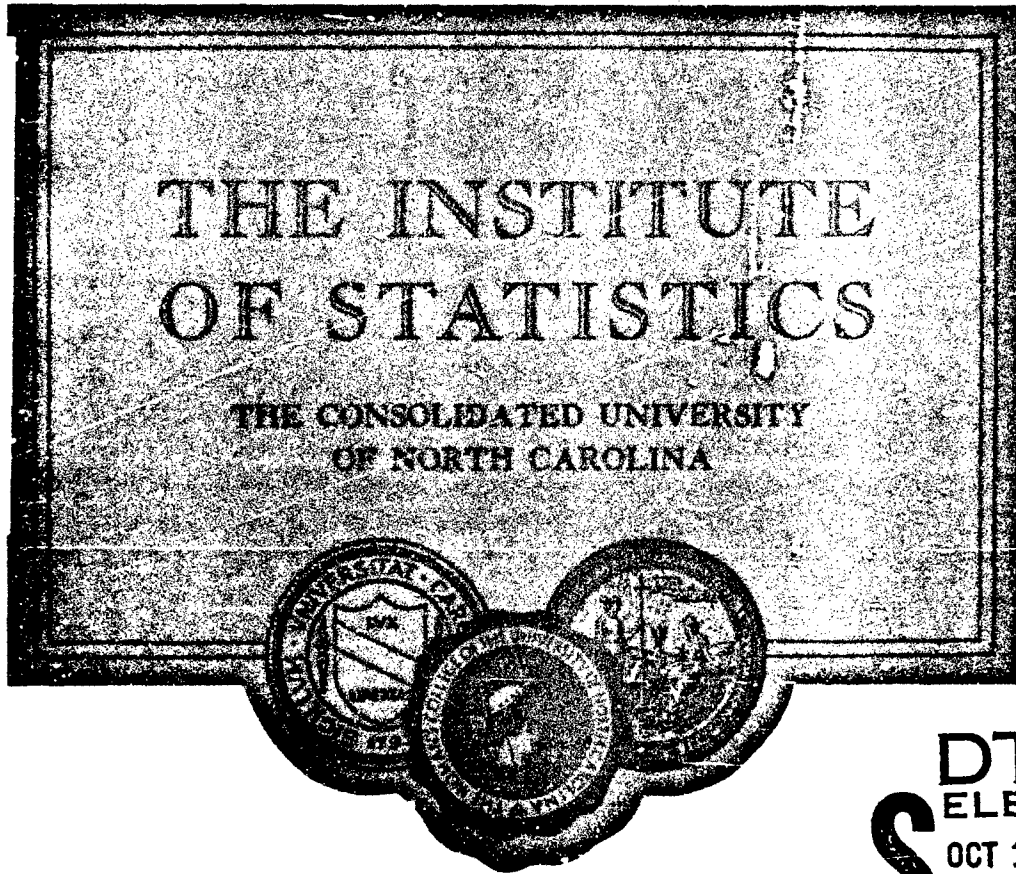


AD-A160 190



DTIC  
ELECTE  
OCT 11 1985  
S B D

THE LIMITING DISTRIBUTION OF LEAST SQUARES  
IN AN ERRORS-IN-VARIABLES LINEAR REGRESSION MODEL

by

Leon Jay Gleser  
Purdue University

Raymond J. Carroll  
University of North Carolina

Paul P. Gallo  
Lederle Laboratories

Mimeo Series #1577

April 1985

Approved for public release;  
distribution unlimited.

DEPARTMENT OF STATISTICS  
Chapel Hill, North Carolina

DTIC FILE COPY

NONCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

*Alto 190*

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Nonclassified		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Mimeo Series #1577		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR. 85-0869	
6a. NAME OF PERFORMING ORGANIZATION University of N.C.	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research	
6c. ADDRESS (City, State and ZIP Code) Department of Statistics, Phillips Hall Chapel Hill, N.C. 27514		7b. ADDRESS (City, State and ZIP Code) <i>Bolling AF BDC 20332</i>	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620 82 C 0009	
8c. ADDRESS (City, State and ZIP Code) Bolling Air Force Base Washington, D.C. 20332		10. SOURCE OF FUNDING NOS.	
		PROGRAM ELEMENT NO. 61102F	TASK NO. A5
		PROJECT NO. 2304	WORK UNIT NO.
11. TITLE (Include Security Classification) "The Limiting Distribution of Least Squares in an Errors-in-Variables Linear Regression Mode			
12. PERSONAL AUTHOR(S) Gleser, Leon Jay; Carroll, Raymond J. and Gallo, Paul, P.			
13a. TYPE OF REPORT Technical	13b. TIME COVERED FROM 9/84 TO 8/85	14. DATE OF REPORT (Yr., Mo., Day) April 1985	15. PAGE COUNT 24
16. SUPPLEMENTARY NOTATION			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB. GR.	
		Consistency, Asymptotic normality, Regression, Functional models, Structural models, Instrumental variables, Ordinary least squares estimators.	
19. ABSTRACT (Continue on reverse if necessary and identify by block number)			
<p>It is well-known that the ordinary least squares (OLS) estimator <math>\hat{\beta}</math> of the slope and intercept parameters <math>\beta</math> in a linear regression model with errors of measurement for some of the independent variables (predictors) is inconsistent. However, Gallo (1982) has shown that certain linear combinations of <math>\beta</math> are consistently estimated by the corresponding linear combinations of <math>\hat{\beta}</math>. In this paper, it is shown that under reasonable regularity conditions such linear combinations are (jointly) asymptotically normally distributed. Some methodological consequences of our results are given in a companion paper (Carroll, Gallo and Gleser, 1985).</p>			
<h2>Best Available Copy</h2>			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS <input type="checkbox"/>		21. ABSTRACT SECURITY CLASSIFICATION	
22a. NAME OF RESPONSIBLE INDIVIDUAL Brian W. Woodruff, MAJ, USAF	22b. TELEPHONE NUMBER (Include Area Code) <i>(202) 767-5026</i>	22c. OFFICE SYMBOL AFOSR/NM	

Best Available Copy

THE LIMITING DISTRIBUTION OF LEAST SQUARES  
IN AN ERRORS-IN-VARIABLES LINEAR REGRESSION MODEL

by

Leon Jay Gleser  
Purdue University

Raymond J. Carroll  
University of North Carolina

Paul P. Gallo  
Lederle Laboratories



**S** DTIC  
ELECTE **D**  
OCT 11 1985

B

Accession For	
NTIS GR&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Special
A-1	

-A-

THE LIMITING DISTRIBUTION OF LEAST SQUARES  
IN AN ERRORS-IN-VARIABLES LINEAR REGRESSION MODEL

BY LEON JAY GLESER<sup>1</sup>, RAYMOND J. CARROLL<sup>2</sup>, AND PAUL P. GALLO  
Purdue University, University of North Carolina, and Lederle Laboratories

It is well-known that the ordinary least squares (OLS) estimator  $\hat{\beta}$  of the slope and intercept parameters  $\beta$  in a linear regression model with errors of measurement for some of the independent variables (predictors) is inconsistent. However, Gallo (1982) has shown that certain linear combinations of  $\beta$  are consistently estimated by the corresponding linear combinations of  $\hat{\beta}$ . In this paper, it is shown that under reasonable regularity conditions such linear combinations are (jointly) asymptotically normally distributed. Some methodological consequences of our results are given in a companion paper (Carroll, Gallo and Gleser, 1985).

---

<sup>1</sup>Research of Leon Jay Gleser supported by NSF Grant DMS-8121948.

<sup>2</sup>Research of Raymond J. Carroll supported by Air Force Office of Scientific Research Contract AFOSR-F-49620-82-C-009 and sponsored by the United States Army under Contract No. DAAG 29-80-C-0041.

AMS (1980) subject classifications. Primary 60F05, 62J05; secondary 62F10, 62H99.

Key words and phrases. Regression, functional models, structural models, instrumental variables, ordinary least squares estimators, consistency, asymptotic normality.

1. Introduction. There is a substantial literature concerning linear regression when some of the predictors (independent variables) are measured with error. Such models are of importance in econometrics (instrumental variables models), psychometrics (correction for attenuation, models of change), and in instrumental calibration studies in medicine and industry. Recent theoretical work concerning maximum likelihood estimation in such models appears in Healy (1980), Fuller (1980), and Anderson (1984), while Reilly and Patino-Leal (1981) take a Bayesian approach.

In an applied context, an investigator may either overlook the measurement errors in the predictors, or choose the classical ordinary least squares (OLS) estimator of the parameters because of its familiarity and ease of use. Certainly, the methodology of classical least squares theory (confidence intervals, multiple comparisons, tests of hypotheses, residual analysis) is considerably more developed than the corresponding errors-in-variables methodology, particularly in samples of moderate size. If the OLS estimator is used, what are the consequences?

Cochran (1968) has given a general discussion of the consequences of using the OLS estimator in errors-in-variables contexts. For the special case of the analysis of covariance (ANCOVA), where the covariates are measured with error, detailed investigations have been done by Lord (1960), De Gracie and Fuller (1972) and Cronbach (1976). It is by now well-known that the OLS estimator  $\hat{\beta}$  of the slope and intercept parameters  $\beta$  in such errors-in-variables models is inconsistent; that is,  $\hat{\beta}$  does not tend in probability to  $\beta$  as the sample size  $n$  becomes infinitely large. However, in ANCOVA with covariates measured with error but balanced (in terms of means) across the design, the OLS estimator of the design effects is known to be consistent. This is shown

in the two-treatment case by Cochran (1968) and DeGracie and Fuller (1980).

More generally, Gallo (1982) has shown that for general linear errors-in-variables regression models, certain linear combinations  $c'\hat{\beta}$  of the OLS estimator are consistent estimators of the corresponding linear combinations of  $\beta$ . Gallo's result is reproduced in Section 2 as Theorem 1.

Let the rows of  $C$  be a basis for all linear combinations  $c'\beta$  of  $\beta$  that are consistently estimated by  $c'\hat{\beta}$ . In the present paper, it is shown that under a reasonable extension of the regularity conditions given by Gallo (1982),  $n^{\frac{1}{2}}(C\hat{\beta}-C\beta)$  has a limiting asymptotic multivariate normal distribution (Theorem 2 of Section 2). This result does not require that the random errors (errors of measurement, residual errors) are normally distributed, but only that these errors are sampled from a common population with finite second moments. However, Theorem 2 does assume that all predictors are fixed. In Section 3, Theorem 2 is extended to cases where some of the predictors are random variables.

The nature of the limiting normal distribution of  $n^{\frac{1}{2}}(C\hat{\beta}-C\beta)$  depends upon whether the predictors measured with error are random (structural errors-in-variables models) or fixed (functional errors-in-variables models). In the former case, the limiting normal distribution has a zero mean vector, while in the latter case the mean vector need not be zero (and is a function of unknown parameters). A companion paper (Carroll, Gallo and Gleser, 1985) uses these results to compare the asymptotic efficiencies of the OLS and maximum likelihood estimators of  $C\beta$  when the errors-in-variables model is of the structural kind.

2. Asymptotic Theory. Suppose that a dependent scalar variable  $y_i$  is related to a vector  $f_{1i}$ :  $p \times 1$  of observable predictors and a vector  $f_{2i}$ :  $q \times 1$  of latent (unobservable) predictors by the model

$$(2.1) \quad y_i = f_{1i}'\beta_1 + f_{2i}'\beta_2 + e_i, \quad i = 1, 2, \dots, n,$$

and that  $f_{2i}$  is observed with error by  $x_i$ , where

$$(2.2) \quad x_i = f_{2i} + u_i, \quad i = 1, 2, \dots, n.$$

For fixed  $(f_{1i}', f_{2i}')$  it is assumed that

$$(2.3) \quad \begin{pmatrix} e_i \\ u_i \end{pmatrix}, \quad 1 \leq i \leq n, \text{ are i.i.d.}$$

with mean vector 0 and covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12}' & \Sigma_{22} \end{pmatrix}; \quad \Sigma_{22}: \quad q \times q.$$

To state the model in vector-matrix form, let

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad F_1 = \begin{pmatrix} f_{11}' \\ \vdots \\ f_{1n}' \end{pmatrix}, \quad F_2 = \begin{pmatrix} f_{21}' \\ \vdots \\ f_{2n}' \end{pmatrix}, \quad X = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix},$$

$$e = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}, \quad U = \begin{pmatrix} u_1' \\ \vdots \\ u_n' \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}.$$

Then

$$(2.4) \quad Y = F_1\beta_1 + F_2\beta_2 + e, \quad X = F_2 + U,$$

where the rows of  $E = (e, U)$  are i.i.d. random vectors with mean vector 0 and covariance matrix  $\Sigma$ .

Note. It is assumed that all design (dummy) variables are included in  $F_1$ . This eliminates the need for separately including an intercept term in the model.

The OLS estimator of  $\beta$  for the model (2.4) is

$$(2.5) \quad \hat{\beta} = \begin{pmatrix} F_1'F_1 & F_1'X \\ X'F_1 & X'X \end{pmatrix}^{-1} \begin{pmatrix} F_1'Y \\ X'Y \end{pmatrix}.$$

2.1 Asymptotic Consistency. To give asymptotic results about  $\hat{\beta}$ , we need to make some assumptions about the sequence

$$(2.6) \quad \underline{f}_n = \{(f_{1i}', f_{2i}') : i = 1, 2, \dots\}$$

of fixed predictor values. These are the following.

Assumption 1.

$$\lim_{n \rightarrow \infty} n^{-1} \begin{bmatrix} F_1'F_1 & F_1'F_2 \\ F_2'F_1 & F_2'F_2 \end{bmatrix} = \begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{12}' & \Delta_{22} \end{bmatrix} \equiv \Lambda, \quad \Lambda > 0.$$



Assumption 2.

$$\lim_{n \rightarrow \infty} n^{-\frac{1}{2}} \max [F_1, F_2] = 0,$$

where for any matrix  $A = ((a_{ij}))$ ,  $\max(A) = \max |a_{ij}|$ .

We will make extensive use of the following results.

Lemma 1. Under (2.4) and Assumptions 1 and 2, for all  $(q+1)$ -dimensional column vectors  $t$ ,

$$n^{-\frac{1}{2}} (F_1, F_2)' (e, U)t \rightarrow \text{MVN}(0, (t' \Sigma t) \Delta)$$

in distribution as  $n \rightarrow \infty$ . In particular,

$$(2.7) \quad n^{-\frac{1}{2}} (f_1, F_2)' (e - U\beta_2) \rightarrow \text{MVN}(0, [(1, -\beta_2') \Sigma \begin{pmatrix} 1 \\ -\beta_2 \end{pmatrix}] \Delta)$$

in distribution as  $n \rightarrow \infty$ .

Proof. This is a direct consequence of Corollary 3.2 and the discussion following in Gleser (1965).  $\square$

Lemma 2. Under the assumptions of Lemma 1,

$$n^{-1} \begin{pmatrix} F_1' F_1 & F_1' X \\ X' F_1 & X' X \end{pmatrix} = \begin{pmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{12}' & \Delta_{22} + \Sigma_{22} \end{pmatrix} + o_p(1).$$

Proof. From the weak law of large numbers,

$$(2.8) \quad n^{-1} (e, U)' (e, U) = \Sigma + o_p(1),$$

while from Lemma 1,

$$n^{-1} F_2' U = o_p(n^{-\frac{1}{2}}).$$

From these facts, (2.4) and Assumption 1, the assertion of the lemma directly follows.  $\square$

The following theorem is a restatement of the result of Gallo (1982) mentioned in Section 1.

Theorem 1 (Gallo, 1982). Under (2.4) and Assumptions 1 and 2,

$$c' \hat{\beta} \xrightarrow{p} c' \beta \Leftrightarrow c' \begin{pmatrix} -\Lambda_{11}^{-1} \Delta_{12} \\ I_q \end{pmatrix} = 0,$$

where  $I_q$  is the  $q$ -dimensional identity matrix.

Proof. Note from (2.4) that

$$\begin{aligned} \frac{1}{n} \left[ \begin{pmatrix} F_1' Y \\ F_2' Y \end{pmatrix} - \begin{pmatrix} F_1' F_1 & F_1' X \\ X' F_1 & X' X \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \right] \\ = \frac{1}{n} \begin{bmatrix} F_1' (e - U\beta_2) \\ F_2' (e - U\beta_2) + U' (e - U\beta_2) \end{bmatrix}. \end{aligned}$$

However, Lemma 1 implies that

$$\frac{1}{n} \begin{pmatrix} F_1' \\ F_2' \end{pmatrix} (e - U\beta_2) = o_p(n^{-\frac{1}{2}}),$$

while it follows from (2.8) that

$$\frac{1}{n} U'(e - U\beta_2) = \sigma_{12}' - \Sigma_{22}\beta_2 + o_p(1).$$

From these facts, (2.5) and Lemma 2 it follows that

$$(2.9) \quad \hat{\beta} = \beta + \begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{12}' & \Delta_{22} + \Sigma_{22} \end{bmatrix}^{-1} \begin{pmatrix} 0 \\ \sigma_{12}' - \Sigma_{22}\beta_2 \end{pmatrix} + o_p(1).$$

Let

$$Q = (\Sigma_{22} + \Lambda_{22.1})^{-1}, \quad \Lambda_{22.1} = \Delta_{22} - \Delta_{12}' \Delta_{11}^{-1} \Delta_{12}.$$

Then

$$\begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{12}' & \Delta_{22} + \Sigma_{22} \end{bmatrix}^{-1} \begin{pmatrix} 0 \\ I_q \end{pmatrix} = \begin{bmatrix} -\Delta_{11}^{-1} \Delta_{12} \\ I_q \end{bmatrix} Q$$

and it follows from (2.9) that

$$(2.10) \quad c' \hat{\beta} \xrightarrow{P} c' \beta + c' \begin{bmatrix} -\Delta_{11}^{-1} & \Delta_{12} \\ I_q & \end{bmatrix} Q(\sigma'_{12} - \Sigma_{22} \beta_2).$$

Thus,

$$c' \hat{\beta} \xrightarrow{P} c' \beta = c' \begin{bmatrix} \Delta_{11}^{-1} & \Delta_{12} \\ I_q & \end{bmatrix} Q(\sigma'_{12} - \Sigma_{22} \beta_2) = 0, \quad \text{all } \beta, \Sigma.$$

Clearly

$$c' \begin{bmatrix} -\Delta_{11}^{-1} & \Delta_{12} \\ I_q & \end{bmatrix} = 0 \Rightarrow c' \begin{bmatrix} -\Delta_{11}^{-1} & \Delta_{12} \\ I_q & \end{bmatrix} Q(\sigma'_{12} - \Sigma_{22} \beta_2) = 0$$

for all  $\beta, \Sigma$ . On the other hand, if

$$\beta_2 = \Sigma_{22}^{-1} \sigma'_{22} - \Sigma_{22}^{-1} (-\Delta_{12}' \Delta_{11}^{-1}, I_q) c,$$

then

$$\begin{aligned} 0 &= c' \begin{bmatrix} -\Delta_{11}^{-1} & \Delta_{12} \\ I_q & \end{bmatrix} Q(\sigma'_{12} - \Sigma_{22} \beta_2) = c' \begin{pmatrix} -\Delta_{11}^{-1} & \Delta_{12} \\ I_q & \end{pmatrix} Q \begin{pmatrix} -\Delta_{11}^{-1} & \Delta_{12} \\ I_q & \end{pmatrix}' c = 0 \\ &= c' \begin{pmatrix} -\Delta_{11}^{-1} & \Delta_{12} \\ I_q & \end{pmatrix} = 0, \end{aligned}$$

since  $Q > 0$ . This completes the proof.  $\square$

Note that

$$c' \begin{bmatrix} -\Delta_{11}^{-1} & \Delta_{12} \\ I_q & \end{bmatrix} = 0 \Leftrightarrow c = d' [I_p, \Delta_{11}^{-1} \Delta_{12}], \text{ some } d.$$

From this fact, it is easily seen that the rows of

$$C = (I_p, \Delta_{11}^{-1} \Delta_{12})$$

serve as a basis for the linear manifold of all  $c$  such that  $c'\beta$  is consistent for  $c'\hat{\beta}$ . This motivates consideration of the limiting distribution of

$$T_n = n^{\frac{1}{2}} C(\hat{\beta} - \beta).$$

2.2 Asymptotic Normality of  $T_n$ . Rather than state our main result (Theorem 2) at once, we first derive a representation for  $T_n$  that leads us to the extra assumption needed to obtain asymptotic normality of  $T_n$ .

Let

$$(L_{1n}, L_{2n}) = C \left[ \frac{1}{n} \begin{pmatrix} F_1' F_1 & F_1' X \\ X' F_1 & X' X \end{pmatrix} \right]^{-1}$$

and

$$\begin{pmatrix} W_{1n} \\ W_{2n} \end{pmatrix} = \frac{1}{n} \left[ \begin{pmatrix} F_1' Y \\ X' Y \end{pmatrix} - \begin{pmatrix} F_1' F_1 & F_1' X \\ X' F_1 & X' X \end{pmatrix} \left( \beta + \begin{pmatrix} -\Delta_{11}^{-1} & \Delta_{12} \\ I_q & \end{pmatrix} \gamma \right) \right]$$

where

$$\gamma = Q(\sigma_{12}^2 - \Sigma_{22} \beta_2).$$

Since

$$C \begin{pmatrix} -\Delta_{11}^{-1} & \Delta_{12} \\ I_q & \end{pmatrix} = 0,$$

it follows from (2.5) that

$$(2.11) \quad T_n = n^{\frac{1}{2}} (L_{1n}, L_{2n}) \begin{pmatrix} W_{1n} \\ W_{2n} \end{pmatrix}.$$

Lemma 3. Under the assumptions of Lemma 1,

$$L_{1n} = \Delta_{11}^{-1} + o_p(1),$$

and

$$G_n = n^{\frac{1}{2}} L_{1n} (W_{1n} + \frac{1}{n} F_1' (F_2 - F_1 \Delta_{11}^{-1} \Delta_{12}) \gamma) \\ \longrightarrow \text{MVN}(0, \left\{ \begin{bmatrix} 1 \\ -(B_2 + \gamma) \end{bmatrix}' \Sigma \begin{bmatrix} 1 \\ -(B_2 + \gamma) \end{bmatrix} \right\} \Delta_{11}^{-1})$$

in distribution as  $n \rightarrow \infty$ .

Proof. The first assertion is a direct consequence of Lemma 2 and the fact that

$$c \begin{pmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{12}' & \Delta_{22} + \Sigma_{22} \end{pmatrix}^{-1} = (\Delta_{11}^{-1}, 0)$$

Note from (2.4) and the definition of  $W_{1n}$  that

$$W_{1n} + \frac{1}{n} F_1'(F_2 - F_1 \Delta_{11}^{-1} \Delta_{12}) \gamma = \frac{1}{n} F_1'(e, U) \begin{pmatrix} 1 \\ -(\beta_2 + \gamma) \end{pmatrix}.$$

The second assertion of the lemma now follows from this representation, Lemma 1, the first assertion of the lemma and Slutsky's Theorem.  $\square$

Lemma 4. Under the assumptions of Lemma 1,

$$(2.12) \quad \begin{aligned} W_{2n} = & \left[ \frac{1}{n} U'(e - U(\beta_2 + \gamma)) - \Delta_{22.1} \gamma \right] \\ & - \left[ \frac{1}{n} F_2'(F_2 - F_1 \Delta_{11}^{-1} \Delta_{12}) - \Delta_{22.1} \right] \gamma + O_p(n^{-\frac{1}{2}}) \end{aligned}$$

and

$$L_{2n} = -\left(\frac{1}{n} F_1' F_1\right)^{-1} \left(\frac{1}{n} F_1'(F_2 - F_1 \Delta_{11}^{-1} \Delta_{12})\right) [Q^{-1} + o_p(1)]^{-1} + O_p(n^{-\frac{1}{2}}).$$

Proof. Using (2.4) and the definition of  $W_{2n}$ , we can write  $W_{2n}$  as the sum of the first two terms on the right-hand side of (2.12) plus

$$\frac{1}{n} F_2'(e - U(\beta_2 + \gamma)) - \frac{1}{n} U'(F_1, F_2) \begin{pmatrix} -\Delta_{11}^{-1} & \Delta_{12} \\ & I_p \end{pmatrix} \gamma.$$

Using Lemma 1, this last term can be shown to be  $O_p(n^{-\frac{1}{2}})$ , as asserted.

From facts about inverses of partitioned matrices, the definitions of  $C$  and  $L_{2n}$  and (2.4),

$$L_{2n} = -\left(\frac{1}{n} F_1' F_1\right)^{-1} \left[\frac{1}{n} F_1' (F_2 - F_1 \Delta_{11}^{-1} \Delta_{12}) + \frac{1}{n} F_1' U\right] A_n$$

where

$$A_n^{-1} = \frac{1}{n} (X'X - X'F_1 (F_1'F_1)^{-1} F_1'X).$$

Using Lemma 2, it is easily shown that

$$A_n^{-1} = \Delta_{22.1} + \Sigma_{22} + o_p(1) = Q^{-1} + o_p(1).$$

Using Lemma 1,

$$n^{-1} F_1' U = o_p(n^{-\frac{1}{2}}).$$

Since  $n^{-1} F_1' F_1 = \Delta_{11} + o(1)$  by Assumption 1, the representation for  $L_{2n}$  given by the lemma follows from Slutsky's Theorem.  $\square$

Using (2.8), Assumption 1 and Lemma 4, it is straightforward to show that  $W_{2n} = o_p(1)$ . Let

$$(2.13) \quad Z_n = n^{-\frac{1}{2}} F_1' (F_2 - F_1 \Delta_{11}^{-1} \Delta_{12}).$$

It follows from (2.11) and Lemmas 3 and 4 that

$$(2.14) \quad T_n = G_n - (\Delta_{11}^{-1} + o_p(1)) Z_n \gamma - (\Delta_{11}^{-1} + o(1)) Z_n [Q^{-1} + o_p(1)]^{-1} (o_p(1)) + o_p(1).$$



A careful look at (2.14) shows that for  $T_n$  to converge in distribution for all  $\beta, \varepsilon$  it is necessary that  $Z_n$  be  $O(1)$ . Thus, we are led to make the following assumption

Assumption 3. For every sequence  $f$  defined by (2.6),

$$\lim_{n \rightarrow \infty} Z_n = \lim_{n \rightarrow \infty} n^{-\frac{1}{2}} F_1'(F_2 - F_1 \Delta_{11}^{-1} \Delta_{12}) = Z(f)$$

where the limit  $Z(f)$  may depend on  $f$ .

That Assumption 3, together with Assumptions 1 and 2, is sufficient for  $T_n$  to have a limiting multivariate normal distribution is clear from (2.13), Lemma 3 and Slutsky's Theorem. This is our main result.

Theorem 2. Under Assumptions 1, 2 and 3,

$$T_n = n^{\frac{1}{2}} (C\hat{\beta} - C\beta) \rightarrow \text{MVN}(-\Delta_{11}^{-1} Z(f)\gamma, (n'\Sigma n)\Delta_{11}^{-1})$$

in distribution as  $n \rightarrow \infty$ , where  $C = (I_p, \Delta_{11}^{-1} \Delta_{12})$ ,

$$\gamma = (\Sigma_{22} + \Delta_{22.1})^{-1} (\alpha_{12}' - \Sigma_{22}\beta_2), \quad n' = (1, -(\beta_2 + \gamma)')$$

3. Discussion and Extensions. Theorems 1 and 2 assume that the sequence  $f$  defined by (2.6) is a sequence of fixed vectors. If elements of the vectors  $(f_{1i}', f_{2i}')$  in this sequence are random variables, one can think of these results as being conditional limit theorems.

When components of each  $(f_{1i}', f_{2i}')$ ,  $i = 1, 2, \dots$ , are random, a fairly easy argument can be used to extend Theorems 1 and 2 to apply unconditionally, provided that  $\Delta_{11}^{-1} Z_n \gamma$ , where  $Z_n = Z_n(f)$  is defined by (2.13), has an asymptotic distribution.

Thus, let  $s_i$  represent the random part of  $(f'_{1i}, f'_{2i})$  and let  $\xi = (s_i, i=1,2,\dots)$ . Distributional assumptions about the  $s_i$  yield a probability measure  $\mu(\xi)$  over the sequences  $\xi$ . Suppose that

$$A = \{\xi: \lim_{n \rightarrow \infty} n^{-1} (F_1, F_2)' (F_1, F_2) = \Delta > 0, \quad \lim_{n \rightarrow \infty} n^{-\frac{1}{2}} (F_1, F_2) = 0\}$$

satisfies

$$(3.1) \quad \int_A d\mu(\xi) = 1.$$

In other words, Assumptions 1 and 2 are satisfied with probability one. Then Theorem 1 shows that for all  $\xi$  in  $A$ , all  $\epsilon > 0$ ,

$$\lim_{n \rightarrow \infty} P\{[\text{tr}(\hat{C}\hat{\beta} - C\beta)'(\hat{C}\hat{\beta} - C\beta)]^{\frac{1}{2}} > \epsilon | \xi\} = 0.$$

Thus, by the Lebesgue Dominated Convergence Theorem, for all  $\epsilon > 0$ ,

$$\lim_{n \rightarrow \infty} P\{[\text{tr}(\hat{C}\hat{\beta} - C\beta)'(\hat{C}\hat{\beta} - C\beta)]^{\frac{1}{2}} > \epsilon\} = 0,$$

and hence  $\hat{C}\hat{\beta}$  converges unconditionally in probability to  $C\beta$ . This shows that Theorem 1 holds unconditionally (over  $\xi$ ).

In a similar fashion, it can be shown that the representation (2.14) for  $T_n$  holds unconditionally, that  $G_n$  in that representation has the limiting multivariate normal distribution described in Lemma 3, and that  $G_n$  and  $Z_n$  are asymptotically statistically independent. Consequently, if  $\Delta_{11}^{-1} Z_n \gamma$  has a limiting distribution, the limiting distribution of  $T_n$  is the convolution of the limiting distributions of  $G_n$  and  $-\Delta_{11}^{-1} Z_n \gamma$ .

Note: The above discussion is only a sketch of the arguments needed, and skips over such details as measurability. A more extensive discussion in a similar context can be found in Gleser (1983).

We will now follow the steps of the above analysis for some special cases of the model (2.4) which are commonly adopted in practice. Recall that if  $f_{2i}$ ,  $i = 1, 2, \dots$ , are random vectors, the model (2.4) is called a structural linear errors-in-variables regression model, while if the  $f_{2i}$ ,  $i = 1, 2, \dots$ , are vectors of constants, the model is that of a functional linear errors-in-variables regression model. Mixes of these cases, where some elements of  $f_{2i}$  are fixed and some elements are random, are also possible. Further, the elements of  $f_{1i}$  (except for the first component, which is always equal to 1 to accommodate an intercept term) can also be fixed or random. Let

$$f_{1i} = \begin{pmatrix} 1 \\ h_i \end{pmatrix}.$$

We will consider the following cases:

- (a) both  $h_i$  and  $f_{2i}$  fixed (functional model),
- (b)  $h_i$  random,  $f_{2i}$  fixed (functional model),
- (c)  $h_i$  fixed,  $f_{2i}$  random (structural model),
- (d) both  $h_i$  and  $f_{2i}$  random (structural model).

3.1 Both  $h_i$  and  $f_{2i}$  fixed. Theorems 1 and 2 already summarize what we can say about this case. Although Theorem 2 has some technical interest, it is unfortunately rather useless for statistical applications. Unless we are in the unlikely case where we either know the limit  $Z(f)$  or can consistently estimate this quantity, we cannot use Theorem 2 to construct large-sample confidence regions

for C8. Recall that  $\{f_{2i}, i = 1, 2, \dots\}$  is a sequence of unknown parameters, and that the individual vectors  $f_{2i}$  in this sequence cannot be consistently estimated. Thus, very strong assumptions are needed to permit us to consistently estimate  $Z(f)$  (or  $\Delta_{11}^{-1} Z(f)\gamma$ ).

3.2.  $h_i$  random and  $f_{2i}$  fixed. Here, we can assume that the vectors  $h_i$  are mutually statistically independent, but must consider the possibility that the distribution of  $h_i$  depends upon  $f_{2i}$ ,  $i = 1, 2, \dots$ . (That is, the  $h_i$ 's are not identically distributed.) Given the linear form of (2.4), it is natural to assume that a similar linear model relates  $h_i$  to  $f_{2i}$ . Thus, we assume that

$$(3.2) \quad h_i = \alpha + \psi f_{2i} + t_i, \quad i = 1, 2, \dots$$

where the  $t_i$ 's are i.i.d. with mean vector 0 and covariance matrix  $\Lambda$ . We also assume that

$$(3.3) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f_{2i} = \mu, \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f_{2i} f_{2i}' = \Delta_{22} > 0$$

and that  $\lim_{n \rightarrow \infty} n^{-\frac{1}{2}} f_{2i} = 0$ , all  $i$ . By letting  $f_{2i} \rightarrow f_{2i} - \mu$ ,  $\alpha \rightarrow \alpha + \psi\mu$ ,

$\Delta_{22} \rightarrow \Delta_{22} - \mu\mu'$ ; we can let  $\mu = 0$  without loss of generality.

The strong law of large numbers shows that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n t_i = 0, \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n t_i t_i' = \Lambda$$

with probability one. Using (3.2), (3.3) and Theorem 3 of Chow (1966),

$$\lim_{n \rightarrow \infty} n^{-\frac{1}{2}} \sum_{i=1}^n t_i f'_{2i} \left( \frac{1}{n} \sum_{j=1}^n f_{2j} f'_{2j} \right)^{-\frac{1}{2}} = 0$$

with probability one. Thus (3.1) holds with

$$\Delta = \begin{pmatrix} 1 & \alpha' & 0 \\ \alpha & \alpha\alpha' + \psi\Delta_{22}\psi' + \Lambda & \psi\Delta_{22} \\ 0 & \Delta_{22}\psi' & \Delta_{22} \end{pmatrix}.$$

Note that

$$\Delta_{11}^{-1} \Delta_{12} = \begin{bmatrix} -\alpha' \\ I_{p-1} \end{bmatrix} [\psi\Delta_{22}\psi' + \Lambda]^{-1} \psi\Delta_{22}.$$

Let  $1'_n = (1, 1, \dots, 1)$  and  $T' = (t_1, \dots, t_n)$ . Then

$$\begin{aligned} Z_n &= n^{-\frac{1}{2}} F'_1 (F_2 - F_1 \Delta_{11}^{-1} \Delta_{12}) \\ &= n^{-\frac{1}{2}} \begin{pmatrix} 1'_n \\ \alpha 1'_n + \psi F'_2 + T' \end{pmatrix} (F_2 \Gamma - T \Omega) \end{aligned}$$

where

$$\Gamma = I_q - \psi' \psi, \quad \Omega = [\psi\Delta_{22}\psi' + \Lambda]^{-1} \psi\Delta_{22}.$$

It is apparent that, in general, extra conditions on both  $F_2$  and the higher order moments of the common distribution of the  $t_i$ 's are needed to permit  $Z_n$  to have a limiting distribution.

However, consider the special case  $\psi = 0$ . In this case the random parts  $h_i$  of  $f_{1i}$  are i.i.d. random vectors independent of the  $f_{2i}$ 's, and

$$\Delta_{11}^{-1} Z_n \gamma = n^{-\frac{1}{2}} \Delta_{11}^{-1} \begin{pmatrix} 1' F_2 \\ \alpha 1' F_2 + T' F_2 \end{pmatrix} \gamma = n^{-\frac{1}{2}} \begin{pmatrix} 1' F_2 \gamma \\ \Lambda^{-1} T' F_2 \gamma \end{pmatrix}$$

Using Corollary 3.2 and the discussion following in Gleser (1965), it can be shown that the elements of  $n^{-\frac{1}{2}} T' F_2 \gamma$  have an asymptotic multivariate normal distribution:

$$n^{-\frac{1}{2}} T' F_2 \gamma \rightarrow \text{MVN}(0, (\gamma' \Delta_{22} \gamma) \Lambda).$$

Although we could impose the condition that  $n^{-\frac{1}{2}} 1' F_2 \gamma = o_p(1)$ , this is a rather restrictive condition, and still leaves us the problem of estimating the limit of  $n^{-\frac{1}{2}} 1' F_2 \gamma$  in statistical applications. Instead, we settle for a more restricted result:

$$(3.4) \quad n^{\frac{1}{2}} (0, I_{p-1})(\hat{C}B - CB) \rightarrow \text{MVN}(0, \Theta)$$

in distribution as  $n \rightarrow \infty$ , where

$$\begin{aligned} \Theta &= \begin{pmatrix} 1 \\ -(\beta_2 + \gamma) \end{pmatrix}' \Sigma \begin{pmatrix} 1 \\ -(\beta_2 + \gamma) \end{pmatrix} (0, I_{p-1}) \Delta_{11}^{-1} \begin{pmatrix} 0 \\ I_{p-1} \end{pmatrix} \\ &\quad + (\gamma' \Delta_{22} \gamma) \Lambda^{-1} \\ &= \Lambda^{-1} \left[ \begin{pmatrix} 1 \\ -(\beta_2 + \gamma) \end{pmatrix}' \Sigma \begin{pmatrix} 1 \\ -(\beta_2 + \gamma) \end{pmatrix} + \gamma' \Delta_{22} \gamma \right], \end{aligned}$$

since  $\Lambda^{-1} = (0, I_{p-1}) \Delta_{11}^{-1} (0, I_{p-1})'$ . In this context ( $\psi=0$ ), it is worth noting that

$$\begin{aligned} (0, I_{p-1}) C &= (0, I_{p-1}) (I_p, -\Delta_{11}^{-1} \Delta_{12}) \\ &= (0, I_{p-1}, 0), \end{aligned}$$

so that the result concerns the estimates of the slopes  $(0, I_{p-1}) \beta_1$  of the  $y_i$  on the  $h_i$  (the random part of  $f_{1i}$ ) in (2.4).

3.3  $h_i$  fixed and  $f_{2i}$  random. In analogy with the discussion in Section 3.2, we assume that

$$(3.5) \quad f_{2i} = \psi f_{1i} + t_i, \quad i = 1, 2, \dots,$$

where the  $t_i$  are i.i.d. with common mean vector 0 and covariance matrix  $\Lambda$ . (Here, since the first element of  $f_{1i}$  is always 1, there is no need for a separate intercept term.) Assumption (3.5) is commonly adopted in instrumental variables approaches to errors in variables models in econometrics, and in ANCOVA with measurement errors in the covariates.

We also assume that

$$(3.6) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f_{1i} f_{1i}' = \Delta_{11} > 0$$

and that  $\lim_{n \rightarrow \infty} n^{-\frac{1}{2}} f_{1i} = 0$ , all  $i$ . Following steps similar to those used in Section 3.2, we can show that (3.1) holds with

$$\Delta = \begin{pmatrix} \Delta_{11} & \Delta_{11}\psi' \\ \psi\Delta_{11} & \psi\Delta_{11}\psi' + \Lambda \end{pmatrix}.$$

Hence,

$$\Lambda_{11}^{-1}\Delta_{12} = \psi'.$$

Note that

$$Z_n = n^{-\frac{1}{2}} F_1'(F_2 - F_1\Delta_{11}^{-1}\Delta_{12}) = n^{-\frac{1}{2}} F_1'T,$$

where  $T' = (t_1, \dots, t_n)$ . Applying Corollary 3.2 and the following discussion in Gleser (1965),

$$\Delta_{11}^{-1}Z_n Y \rightarrow \text{MVN}(0, \Delta_{11}^{-1}(\gamma' \Lambda \gamma))$$

in distribution as  $n \rightarrow \infty$ . Consequently,

$$(3.7) \quad n^{\frac{1}{2}}(C\hat{\beta} - C\beta) \rightarrow \text{MVN}(0, \Delta_{11}^{-1} [n'\Sigma n + \gamma' \Lambda \gamma])$$

in distribution as  $n \rightarrow \infty$ . It is worth noting that here

$$C = (I_p, \psi), \quad \Lambda = \Delta_{22.1}, \quad \eta = \begin{pmatrix} 1 \\ -(\beta_2 + \gamma) \end{pmatrix}.$$

When  $\psi = 0$ , there is a close parallel between (3.4) and (3.7). Note also that in this case  $C\beta = \beta_1$ .



Even when  $\psi \neq 0$  (the distribution of  $f_{2i}$  depends on  $f_{1i}$ ), the result (3.7) was obtained without the need to make extra assumptions on the higher moments of the common distribution of the  $t_i$ , in contrast to our conclusions in the case of Section 3.2.

3.4 Both  $h_i$  and  $f_{2i}$  random. In this case it is more natural to make assumptions concerning  $(h_i', f_{2i}')$ ,  $i = 1, 2, \dots$ . We assume that these vectors are i.i.d. with a common mean vector  $\mu$  and a common covariance matrix  $\phi$ . The strong law of large numbers now shows that (3.1) holds with

$$\Delta = \begin{pmatrix} 1 & \mu' \\ \mu & \phi + \mu\mu' \end{pmatrix}.$$

Let  $\mu' = (\mu_1', \mu_2')$  and

$$\phi = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{12}' & \phi_{22} \end{pmatrix}$$

where  $\mu_1'$ ,  $\phi_{11}$  are the common mean vector and covariance matrix of the  $h_i$ 's.

Thus,

$$\begin{aligned} \phi_{11}^{-1} \phi_{12}' &= \begin{pmatrix} 1 & \mu_1' \\ \mu_1 & \phi_{11} + \mu_1 \mu_1' \end{pmatrix}^{-1} \begin{pmatrix} \mu_2' \\ \phi_{12} + \mu_1 \mu_2' \end{pmatrix} \\ &= \begin{pmatrix} \mu_2' - \mu_1' \phi_{11}^{-1} \phi_{12}' \\ \phi_{11}^{-1} \phi_{12}' \end{pmatrix}. \end{aligned}$$

(Best Avail. Copy)

Let  $H' = (h_1, h_2, \dots, h_n)$ . Then

$$Z_n = n^{-\frac{1}{2}} \begin{pmatrix} 1' \\ n \\ H' \end{pmatrix} (F_2 - 1_n(\mu_2' - \mu_1' \phi_{11}^{-1} \phi_{12}) - H \phi_{11}^{-1} \phi_{12}).$$

The Central Limit Theorem shows that the first row of  $Z_n$  has an asymptotic multivariate normal distribution. For the remaining rows of  $Z_n$  to be asymptotically multivariate normally distributed, additional assumptions on the higher moments of the joint distribution of  $(h_i', f_{2i}')$  are needed. To avoid such assumptions, we can assume that

$$(3.8) \quad f_{2i} = \mu_2 - \phi_{12}' \phi_{11}^{-1} \mu_1 + \phi_{12}' \phi_{11}^{-1} h_i + t_i, \quad i = 1, 2, \dots,$$

where the  $t_i$ 's are i.i.d. with mean vector 0 and covariance matrix

$$\phi_{22.1} = \phi_{22} - \phi_{12}' \phi_{11}^{-1} \phi_{12}$$

and statistically independent of the  $h_i$ 's. If we condition on the  $h_i$ 's, (3.8) is the model (3.5) with

$$\psi = (\mu_2 - \phi_{12}' \phi_{11}^{-1} \mu_1, \phi_{12}' \phi_{11}^{-1}), \quad \Lambda = \phi_{22.1}.$$

We can now use the results of Section 3.2, noting that with probability one (over sequences  $h_1, h_2, \dots$ )

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} F_1' F_1 &= \lim_{n \rightarrow \infty} \frac{1}{n} (1_n, H)' (1_n, H) \\ &= \begin{pmatrix} 1 & \mu_j' \\ \mu_j & \phi_{11} + \mu_j \mu_j' \end{pmatrix} = \Delta_{11}. \end{aligned}$$

Thus, conditional on the  $h_j$ 's,

$$(3.9) \quad n^{\frac{1}{2}} (\hat{C}\beta - C\beta) \rightarrow \text{MVN}(0, \Delta_{11}^{-1} [n' \Sigma n + \gamma' \phi_{22} \gamma])$$

in distribution as  $n \rightarrow \infty$ . Using the arguments given at the beginning of this section about converting conditional limiting results to unconditional limiting results, we conclude that (3.9) also holds unconditionally.

**3.5 Conclusion.** The results (3.4), (3.7), (3.9) can be used to construct large sample confidence ellipsoids for  $C\beta$  based on the OLS estimator  $\hat{C}\beta$  provided that consistent estimators can be found for the covariance matrices of the asymptotic normal distributions. It should be noted that in general  $C\beta$  is a function not only of  $\beta$ , but also of  $\Delta_{11}^{-1} \Delta_{12}$ , which need not be a known matrix.

# Best Available Copy

## REFERENCES

- ANDERSON, T. W. (1984). Estimating linear statistical relationships. Ann. Statist. 12, 1-45.
- CARROLL, R. J., GALLO, P. and GLESER, L. J. (1985). Comparisons of least squares and errors-in-variables regression with special reference to randomized analysis of covariance. J. Amer. Statist. Assoc. To appear.
- CHOW, Y. S. (1966). Some convergence theorems for independent random variables. Ann. Math. Statist. 37, 1482-1493.
- CRONBACH, L. J. (1976). On the design of educational measures. In Advancement in Psychological and Educational Measurement, ed. D.N.M. DeGruijter and L. J. Th. van der Kamp, John Wiley and Sons, New York.
- DeGRACIE, J. S. and FULLER, W. A. (1980). Estimation of the slope and analysis of covariance when the concomitant variable is measured with error. J. Amer. Statist. Assoc. 67, 930-937.
- FULLER, W. A. (1980). Properties of some estimators for the errors-in-variables model. Ann. Statist. 8, 407-422.
- GALLO, P. P. (1982). Properties of estimators in errors-in-variables regression models. Ph.D. dissertation, University of North Carolina at Chapel Hill.
- GLESER, L. J. (1965). On the asymptotic theory of fixed-size sequential confidence bounds for linear regression parameters. Ann. Math. Statist. 36, 463-467.
- GLESER, L. J. (1983). Functional, structural and ultrastructural errors-in-variables models. In 1983 Proceedings Bus. and Econ. Statist. Sect., American Statistical Association, Washington, D.C., 57-66.
- HEALY, J. D. (1980). Maximum likelihood estimation of a multivariate linear functional relationship. J. Multivariate Anal. 10, 243-251.
- LORD, F. M. (1960). Large-sample covariance analysis when the control variable is fallible. J. Amer. Statist. Assoc. 55, 307-321.
- REILLY, P. M. and PATINO-LEAL, H. (1981). A Bayesian study of the errors-in-variables model. Technometrics 23, 221-231.

LEON JAY GLESER  
 Department of Statistics  
 Purdue University  
 West Lafayette, Indiana 47907

RAYMOND J. CARROLL  
 Department of Statistics  
 The University of North Carolina  
 321 Phillips Hall 039A  
 Chapel Hill, North Carolina 27514

PAUL P. GALLO  
 Group Leader, Preclinical Statistics  
 Lederle Laboratories  
 Pearl River, New York 10954