The Annals of Statistics
1984, Vol. 12, No. 1, 87–100

THE INFLUENCE FUNCTION IN THE ERRORS IN VARIABLES PROBLEM[1]

BY GABRIELLE KELLY[2]

Stanford University

This paper focuses on two aspects of the errors in variables problem—variance estimation of the classical estimators of slope and intercept, and the detection of influential observations. The behaviour of the jackknife, bootstrap, normal theory and influence function estimators of variability is examined under a number of sampling situations by Monte Carlo methods. In the multivariate case, perturbation analysis is used to calculate the influence function of the estimator of Gleser (1981). The connection to estimation in linear regression models is discussed. The role of the influence function in the detection of influential observations is considered and an illustration is given by a numerical example.

1. Introduction. In the univariate structural equations model n random vectors $\mathbf{z}_i = (x_i, y_i)^T$ are observed. It is assumed that for each $i = 1, \cdots, n$,

$$\mathbf{z}_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} u_{1i} \\ u_{2i} \end{pmatrix} + \begin{pmatrix} e_{1i} \\ e_{2i} \end{pmatrix} = \mathbf{u}_i + \mathbf{e}_i, \tag{1.1}$$

where

$$u_{2i} = \alpha + \beta u_{1i} \tag{1.2}$$

and the $\mathbf{u}_i$ are independently and identically distributed (i.i.d.) with mean vector $\boldsymbol{\mu}$ and covariance matrix $\Sigma_{\mathbf{U}}$ given by

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \text{ and } \Sigma_{\mathbf{U}} = \begin{pmatrix} \sigma^2_{U_1} & \beta\sigma^2_{U_1} \\ & \beta^2\sigma^2_{U_1} \end{pmatrix}. \tag{1.3}$$

The $\mathbf{e}_i$ are i.i.d. with mean vector $\mathbf{O}$ and covariance matrix given by

$$\Sigma_{\mathbf{E}} = \begin{pmatrix} \sigma^2_1 & o \\ & \sigma^2_2 \end{pmatrix}. \tag{1.4}$$

It is also assumed that for each i,

$$\mathbf{u}_i \text{ and } \mathbf{e}_i \text{ are independent.} \tag{1.5}$$

Let F denote the common distribution function of the $\mathbf{z}_i$. By (1.1)–(1.5) the

Received October 1981; revised August 1983.

[1] Adapted from the author's doctoral dissertation at Stanford University while under the guidance of Rupert G. Miller. Research was supported by NIH Grant R01 GM21215.

[2] Now at University College, Cork, Ireland.

AMS 1970 *subject classifications.* Primary G2E20; secondary 62H10, 62E25, 62J05.

Key words and phrases. Errors in variables, multivariate structural equations model, jackknife, influence function, bootstrap, method of moments, linear regression, Monte Carlo, outliers, robust estimation.

mean vector μ_F and covariance matrix Σ_F are given by

$$\mu(F) = \begin{pmatrix} \mu_X(F) \\ \mu_Y(F) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \alpha + \beta\mu_1 \end{pmatrix}, \tag{1.6}$$

and

$$\Sigma(F) = \begin{pmatrix} \sigma_X^2(F) & \sigma_{XY}(F) \\ & \sigma_Y^2(F) \end{pmatrix} = \begin{pmatrix} \sigma_{U_1}^2 + \sigma_1^2 & \beta\sigma_{U_1}^2 \\ & \beta^2\sigma_{U_1}^2 + \sigma_2^2 \end{pmatrix}. \tag{1.7}$$

Expressions (1.1)–(1.5) represent the model.

When all parameters and all distributions in the model are specified, it is called a structure. A structure generates one and only one distribution $F(\mathbf{z})$ of the observed variables. On the other hand, several structures may generate the same distribution $F(\mathbf{z})$. A parameter is identifiable if it can be uniquely determined from a knowledge of the joint probability distribution of the observed variables. Equivalently, a parameter is identifiable if it has a consistent estimator. For a succinct discussion of identifiability in structural equation models see Riersøl (1950).

It is usually assumed that $\mathbf{U}$ and $\mathbf{E}$ are normally distributed. In order to make the vector of unknown parameters $\boldsymbol{\theta} = (\alpha, \beta)^T$ identifiable in this case, it is necessary to make an additional assumption. The classical assumption is that the ratio

$$\sigma_2^2/\sigma_1^2 = \lambda \tag{1.8}$$

is known. Then, the usual estimator of $\boldsymbol{\theta}$ is the normal theory maximum likelihood estimator (m.l.e.).

In Section 2 of this paper the usual maximum likelihood estimator is viewed as a method of moments estimator. It is seen to be consistent in cases where $\mathbf{U}$ and $\mathbf{E}$ are not normally distributed. The influence function is calculated and used to derive the asymptotic covariance matrix of the estimator. No assumptions on the distribution function F other than the existence of fourth moments are made. The special case where F is a bivariate normal is outlined.

In Section 3 the influence function is used to estimate the asymptotic covariance matrix. This estimator is related to the jackknife. It is seen to be "robust" against departures from assumption (1.4). The connection with estimation in the linear regression model is also considered.

In Section 4, Cook's (1977) measure of influential cases in linear regression is expressed in terms of the influence function. Thus, it has a natural analogue in the errors in variables model and an alternative form is also presented.

Section 5 looks at the multivariate structural equations model. The influence function of the estimator of Gleser (1981) is calculated and used to derive an explicit expression for the asymptotic covariance matrix of the estimators.

In Section 6, the influence function, jackknife, normal theory and bootstrap estimators of the covariance matrix of the m.l.e in the univariate model are investigated by Monte Carlo methods. It is seen that the bootstrap performs well in a variety of sampling situations while the influence function estimator can be considerably biased downward. A brief illustration of the use of the influence function in detecting influential observations is given by a numerical example.

The structural equations model is also referred to as the "errors in variables" problem or the "total least squares" problem. For a review, the reader is referred to Kendall and Stuart (1961, pages 377–382) and Madansky (1959).

2. 2.1 *The method of moments estimator.* Throughout this section we assume that (1.1)–(1.7) and (1.8) hold. The parameter vector $\boldsymbol{\theta} = (\alpha, \beta)^T$ can be written as a functional of the unknown distribution function F by letting

$$\alpha = \alpha(F) = \mu_y(F) - \beta(F)\mu_x(F), \tag{2.1}$$

$$\begin{aligned}\beta &= \beta(F) \\ &= \frac{1}{2\sigma_{xy}(F)}\,[\sigma_y^2(F) - \lambda\,\sigma_x^2(F) + \{(\sigma_y^2(F) - \lambda\,\sigma_x^2(F))^2 + 4\lambda\,\sigma_{xy}^2(F)\}^{1/2}].\end{aligned} \tag{2.2}$$

Let F_n denote the sample distribution function corresponding to F. Denote the sample mean and covariance matrix respectively by

$$\boldsymbol{\mu}(F_n) = \begin{pmatrix}\bar{x}\\ \bar{y}\end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma}(F_n) = \begin{pmatrix}S_X^2 & S_{XY}\\ & S_Y^2\end{pmatrix}. \tag{2.3}$$

Under assumption (1.8) the method of moments estimator of $\boldsymbol{\theta}(F)$ is $\boldsymbol{\theta}(F_n) = (\alpha(F_n), \beta(F_n))^T$ where

$$\alpha(F_n) = \bar{y} - \beta(F_n)\bar{x}, \tag{2.4}$$

$$\beta(F_n) = [S_Y^2 - \lambda S_X^2 + \{(S_Y^2 - \lambda S_X^2)^2 + 4\lambda S_{XY}^2\}^{1/2}]/2S_{XY}. \tag{2.5}$$

By the law of large numbers this estimator is consistent for all distribution functions F with finite second moments.

The role of λ will not be discussed in this paper. Note, however, that a wrong choice of λ will result in $\beta(F_n)$ being inconsistent.

Throughout the remainder of this paper $\boldsymbol{\theta}(F_n) = (\alpha(F_n), \beta(F_n))^T$ will also be denoted by $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\beta})^T$.

When F is the bivariate normal, $\boldsymbol{\theta}(F_n)$ is also the m.l.e.; see Kendall and Stuart (1961, page 381).

Using the definition of influence function, the influence function of $\boldsymbol{\theta}(F)$, $IC(\boldsymbol{\theta}, F; \cdot) = [IC(\alpha, F; \cdot), IC(\beta, F; \cdot)]^T$ can be calculated directly. We have

$$\begin{aligned}IC(\beta, F; \mathbf{z}) = \frac{\beta}{\sigma_{XY}(F)(\beta^2 + \lambda)}\,&\{\beta[(y - \mu_Y(F))^2 - \lambda(x - \mu_X(F))^2] \\ &- (\beta^2 - \lambda)(x - \mu_X(F))(y - \mu_Y(F))\}\end{aligned} \tag{2.6}$$

and

$$IC(\alpha, F; \mathbf{z}) = (y - \alpha - \beta x) - \mu_X(F) IC(\beta, F; \mathbf{z}), \tag{2.7}$$

where $\mathbf{z} = (x, y)^T$. It is easy to check that

$$E_F[IC(\boldsymbol{\theta}, F; \mathbf{z})] = \mathbf{0} \tag{2.8}$$

as required by the definition of the influence function. The sample influence function is obtained by substituting F_n for F i.e. sample moments for population moments, in expressions (2.6) and (2.7).

2.2 *Calculation of the asymptotic covariance matrix.* By the results of Huber (1977, pages 20–22),

$$\sqrt{n}(\theta(F_n) - \theta(F)) \overset{a}{\sim} N_2(\mathbf{0}, E_F[IC(\theta, F; \mathbf{z})IC^T(\theta, F; \mathbf{z})]), \tag{2.9}$$

provided the elements of $E_F[IC(\theta, F; \mathbf{z})IC^T(\theta, F; \mathbf{z})]$ are finite and F is sufficiently smooth. Note that "**a**" denotes "is asymptotically distributed as" when $n \to \infty$.

Squaring and taking the expectation of (2.6) and (2.7) respectively, gives

$$\begin{aligned} &E_F[IC^2(\beta, F; \mathbf{z})] \\ &\quad = \frac{\beta^2}{\sigma_{XY}^2(\beta^2 + \lambda)^2}\{\beta^2[v_{04} + \lambda^2 v_{40}] + (\beta^4 - 4\lambda\beta^2 + \lambda^2)v_{22} \\ &\qquad - 2\beta(\beta^2 - \lambda)[v_{13} - \lambda v_{31}]\} \end{aligned} \tag{2.10}$$

and

$$\begin{aligned} E_F[IC^2(\alpha, F; \mathbf{z})] = \{&v_{02} + \beta^2 v_{20} - 2\beta v_{11} + \mu_x^2 E_F[IC^2(\beta, F; \mathbf{z})] \\ &- 2\mu_X(F)E_F[(y - \alpha - \beta x)IC(\beta, F; \mathbf{z})]\}. \end{aligned} \tag{2.11}$$

For the last term in (2.11)

$$\begin{aligned} &E_F[(y - \alpha - \beta x)IC(\beta, F; \mathbf{z})] \\ &\quad = E_F\{[(y - \mu_Y(F)) - \beta(x - \mu_X(F))]IC(\beta, F; \mathbf{z})\}, \\ &\quad = \frac{\beta}{\sigma_{XY}(\beta^2 + \lambda)} \\ &\qquad \cdot \{\beta v_{03} - \lambda\beta v_{21} - (\beta^2 - \lambda)v_{12} + \lambda\beta^2 v_{30} - \beta^2 v_{12} + \beta(\beta^2 - \lambda)v_{21}\}, \end{aligned} \tag{2.12}$$

where $v_{rs} = E_F\{(x - \mu_x)^r(y - \mu_y)^s\}$.

Also,

$$\begin{aligned} &E_F[IC(\alpha, F; \mathbf{z})IC(\beta, F; \mathbf{z})] \\ &\qquad = E_F[(y - \alpha - \beta x)IC(\beta, F; \mathbf{z})] - \mu_X E_F[IC^2(\beta, F; \mathbf{z})]. \end{aligned} \tag{2.13}$$

Expressions (2.10)–(2.13) determine the entries of $E_F[IC(\theta, F; \mathbf{z})IC^T(\theta, F; \mathbf{z})]$, the asymptotic covariance matrix of $\sqrt{n}\,\theta(F_n)$.

Note that by (2.10)–(2.13) the condition for asymptotic normality (2.9) is equivalent to F having finite fourth moments.

By further algebraic manipulation expressions (2.10) and (2.11) can be written as

$$\begin{aligned} &E_F[IC^2(\beta, F; \mathbf{z})] \\ &\quad = (\beta^2 + \lambda)\sigma_1^2/\sigma_{U_1}^2 + \lambda(\sigma_1^2/\sigma_{U_1}^2)^2(\beta^4 - 4\lambda\beta^2 + \lambda^2)/(\beta^2 + \lambda)^2 \\ &\qquad + (\beta/(\beta^2 + \lambda))^2[(E_F e_2^4 + \lambda^2 E_F e_1^4)/(\sigma_{U_1}^2)^2], \end{aligned} \tag{2.14}$$

and

$$E_F[IC^2(\alpha(F), \mathbf{z})]$$
$$= \sigma_1^2(\beta^2 + \lambda) - 2\mu_1(\beta E_F e_2^3 + \lambda\beta^2 E_F e_1^3)/[\sigma_{U_1}^2(\beta^2 + \lambda)] \tag{2.15}$$
$$+ \mu_1^2 E_F[IC^2(\beta, F; \mathbf{z})]$$

respectively. Thus, we see the asymptotic variability of $\beta(F_n)$ is an increasing function of $\sigma_1^2/\sigma_{U_1}^2$ and $(E_F e_2^4 + \lambda^2 E_F e_1^4)/(\sigma_{U_1}^2)^2$. This makes sense intuitively. As the error variances increase the asymptotic variability of $\beta(F_n)$ increases. Because $\beta(F_n)$ estimates the slope in the linear relationship $U_2 = \alpha + \beta U_1$, as the variance of U_1 gets large the asymptotic variance of $\beta(F_n)$ decreases. Moreover, for fixed σ_1^2, $\sigma_{U_1}^2$, (2.14) implies that the asymptotic variance of $\beta(F_n)$ is minimized when both error distributions have kurtosis equal to zero. This is true when the errors are normally distributed.

For the special case of the bivariate normal, its moment identities can be used to write expressions (2.14) and (2.15) as

$$E_F[IC^2(\beta, F; \mathbf{z})] = \sigma_1^2(\beta^2 + \lambda)/\sigma_{U_1}^2 + \lambda(\sigma_1^2/\sigma_{U_1}^2)^2 \tag{2.16}$$

and

$$E_F[IC^2(\alpha, F; \mathbf{z})] = \sigma_1^2(\beta^2 + \lambda) + \mu_1^2 E_F[IC^2(\beta, F; \mathbf{z})] \tag{2.17}$$

respectively. Here we see the asymptotic variance of $\beta(F_n)$ depends only on the second moments of the underlying distribution function F and is an increasing function of $\sigma_1^2/\sigma_{U_1}^2$

3. Estimation of the asymptotic covariance matrix. We consider three nonparametric estimators of the covariance matrix of $\sqrt{n}\ \theta(F_n)$: the jackknife, the bootstrap and the influence function method. The latter is identical to the infinitesimal jackknife and to the delta method (Efron, 1981). The influence function estimator $\hat{\Sigma}_I$ of the asymptotic covariance matrix $E_F[IC(\theta, F; \mathbf{z})IC^T(\theta, F; \mathbf{z})]$ is defined by

$$\hat{\Sigma}_I = E_{F_n}[IC(\theta, F_n; \mathbf{z})IC^T(\theta, F_n, \mathbf{z})]$$
$$= \frac{1}{n}\sum_{i=1}^n [IC(\theta, F_n; \mathbf{z}_i)IC^T(\theta, F_n; \mathbf{z}_i)]. \tag{3.1}$$

The jackknife estimator $\hat{\Sigma}_J$ is

$$\hat{\Sigma}_J = (n-1)\sum_{i=1}^n [\theta_{-i}(F_n) - \theta_{(\cdot)}(F_n)][\theta_{-i}(F_n) - \theta_{(\cdot)}(F_n)], \tag{3.2}$$

where $\theta_{-i}(F_n)$ is $\theta(F_{n-1})$ with the ith observation omitted and $\theta_{(\cdot)}(F_n) = \sum_{i=1}^n \theta_{-i}(F_n)/n$. The bootstrap estimator $\hat{\Sigma}_B$ must be calculated by a computer algorithm as described in Efron (1981). In the same paper Efron shows how the three methods—jackknife, influence function and bootstrap—derive from the same basic idea and that they are asymptotically equivalent. Thus is suffices to examine the asymptotic properties of any one of them. The asymptotic validity of the bootstrap estimator of standard deviation, for example, is verified in Efron (1979, Section 8, remark G).

Another consistent estimator $\hat{\Sigma}_2$ of the asymptotic covariance can be gotten when F is the bivariate normal, by plugging in the estimates S_X^2, S_Y^2, S_{XY}, $\hat{\beta}$ etc., (i.e., replace F by F_n) in expressions (2.16) and (2.17). Two estimators $\hat{\Sigma}_1$ and $\hat{\Sigma}_2$ arise in the normal case because here one can estimate $E_F(y - \mu_Y)^4$ consistently by either $\sum_{i=1}^n (y_i - \bar{y})^4/n$ or $3[\sum_{i=1}^n (y_i - \bar{y})^2/n]^2$.

The three estimators $\hat{\Sigma}_1$, $\hat{\Sigma}_J$ and $\hat{\Sigma}_B$ are distribution-free not requiring a normal error distribution for their approximate validity, while $\hat{\Sigma}_2$ is not. They also have the following advantage over $\hat{\Sigma}_2$. Suppose for each $i = 1, \cdots, n$

$$\text{Var}(e_{1i}) = \sigma_{1i}^2, \quad \text{Var}(e_{2i}) = \sigma_{2i}^2, \quad \sigma_{2i}^2/\sigma_{1i}^2 = \lambda, \tag{3.3}$$

i.e., the error variances are nonhomogeneous. Conditions (1.1)–(1.3) and (1.5) are retained as before. Then using Taylor series expansions the following lemmas are easily derived, the details of which are omitted here.

LEMMA 1. *$\theta(F_n)$ is consistent under assumption* (3.3).

LEMMA 2. *$\hat{\Sigma}_1$, $\hat{\Sigma}_J$ approximate the true covariance of $\sqrt{n}\,\theta(F_n)$ asymptotically, under assumption* (3.3), *while $\hat{\Sigma}_2$ does not.*

For $\lambda = 0$, $\lambda = +\infty$ the model (1.1)–(1.5) reduces to the regression model of X on Y, Y on X respectively. Taking the case $\lambda = +\infty$ and using (2.6), (2.7), (3.1) and (3.2), expressions for the regression influence function $IC(\theta, F(\lambda = +\infty); \mathbf{z})$ and the estimators of the regression covariance matrix $\hat{\Sigma}_1(\lambda = +\infty)$, $\hat{\Sigma}_2(\lambda = +\infty)$ are easily derived. These are given in Hinkley (1977). Noting that $\{n/(n-2)\}\hat{\Sigma}_2(\lambda = +\infty)$ is the usual regression covariance matrix estimator, Lemma 2 coincides with the results of Hinkley (1977) in the regression case.

4. Detection of influential observations. A general definition of influence is given in Cook and Weisberg (1982, Chapter 3). Here we confine our study to the effect on conclusions when the data is modified by deletion of cases. Cook (1977), Cook and Weisberg (1982) and Atkinson (1982) have discussed the use of $\hat{\theta} - \hat{\theta}_{-i}$ in exhibiting points with a large influence on estimated regression parameters. They all give versions of the statistic D_i of Cook (1977),

$$D_i = (\hat{\theta} - \hat{\theta}_{-i})^T(\hat{\Sigma}_2/n)^{-1}(\hat{\theta} - \hat{\theta}_{-i})/p \tag{4.1}$$

where p is the dimension of θ. Cook noted that a $(1 - \alpha)$ 100% confidence ellipsoid for θ based on $\hat{\theta}$ is given by the set of all θ^* such that

$$(\theta^* - \hat{\theta})^T(\hat{\Sigma}_2/n)^{-1}(\theta^* - \hat{\theta})/p \le F(p, n-p; 1-\alpha) \tag{4.2}$$

where $F(p, n - p; 1 - \alpha)$ is the $(1 - \alpha)$ probability point of the central F distribution with p and $n - p$ degrees of freedom. Thus values of D_i can be compared to the $F(p, n - p)$ distribution. For example if D_i equals the .50 value of the corresponding F distribution, then deletion of the ith case would move the estimate of θ to the edge of a 50% confidence ellipsoid relative to $\hat{\theta}$. An analogy to this can be made in the errors in variables problem. We note that $\hat{\theta}$ is the m.l.e. in the normal case and so asymptotic confidence ellipsoids are given by

$$\{\theta^*: (\theta^* - \hat{\theta})^T(\hat{\Sigma}_2/n)^{-1}(\theta^* - \hat{\theta}) \le \chi_p^2\} \tag{4.3}$$

where χ_p^2 is the chi-squared distribution with p degrees of freedom. Thus, if we here define D_i by

$$D_i = (\hat{\theta} - \hat{\theta}_{-i})^T(\hat{\Sigma}_2/n)^{-1}(\hat{\theta} - \hat{\theta}_{-i}) \tag{4.4}$$

it can be compared to the χ_p^2 distribution. Since $(n-1)(\hat{\theta} - \hat{\theta}_{-i})$ is the jackknife estimate of the influence function (Efron, 1981), an equivalent measure to D_i is V_i, where

$$V_i = IC(\theta, F_n; \mathbf{z}_i)^T(\hat{\Sigma}_1/n)^{-1}IC(\theta, F_n; \mathbf{z}_i)/n^2. \tag{4.5}$$

This is given in Cook and Weisberg (Chapter 3, equation 3.5.25) for the regression case. Of course $\hat{\Sigma}_J$ could also be used instead of $\hat{\Sigma}_2$ in the formula for D_i. We denote the corresponding measure by J_i. The performance of J_i and V_i is examined in the numerical example of Section 5.

Note that the influence function given by (2.6) and (2.7) is unbounded in x and in y respectively. Robust estimation could be used to provide estimators with bounded influence. Robust regression estimators have natural extensions here and will be considered in a later paper.

5. 5.1 *Multivariate structural equations model.* The model given by (1.1)–(1.5) has an extension to the multivariate case and the multivariate analogues of the expressions of Sections 2–4 can be derived.

Let n random vectors $\mathbf{x}_i = (\mathbf{x}_{1i}, \mathbf{x}_{2i})^T$ be observed, where $\mathbf{x}_{1i}$ is $p \times 1$ and $\mathbf{x}_{2i}$ is $r \times 1$, $i = 1, 2, \cdots, n$. It is assumed for each $i = 1, 2, \cdots, n$

$$\mathbf{x}_i = \begin{pmatrix}\mathbf{x}_{1i}\\ \mathbf{x}_{2i}\end{pmatrix} = \begin{pmatrix}\mathbf{u}_{1i}\\ \mathbf{u}_{2i}\end{pmatrix} + \begin{pmatrix}\mathbf{e}_{1i}\\ \mathbf{e}_{2i}\end{pmatrix} = \mathbf{u}_i + \mathbf{e}_i \tag{5.1}$$

where

$$\mathbf{u}_{2i} = \alpha + \mathbf{B}\mathbf{u}_{1i} \tag{5.2}$$

and the $\mathbf{u}_i$ are i.i.d. with mean vector μ and covariance matrix Σ_{U} given by

$$\mu = \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} \quad \text{and} \quad \Sigma_{\mathrm{U}} = \begin{pmatrix}\Sigma_{\mathrm{U}_1} & \mathbf{B}\Sigma_{\mathrm{U}_1}\\ & \mathbf{B}\Sigma_{\mathrm{U}_1}\mathbf{B}^T\end{pmatrix}. \tag{5.3}$$

The $\mathbf{e}_i$ are i.i.d. with mean vector and covariance matrix given by

$$\mathbf{0} \text{ and } \Sigma_{\mathrm{E}} = \sigma^2\mathbf{I}_{p+r}. \tag{5.4}$$

We assume that the

$$\mathbf{u}_i \text{ are independent of the } \mathbf{e}_i. \tag{5.5}$$

Let F be the distribution function of $\mathbf{x}_i$, $i = 1, \cdots, n$ with mean vector $\mu_X(F)$ and covariance matrix $\Sigma(F)$. Let $d_1 > \cdots > d_p > d_{p+1} > \cdots > d_{p+r}$ be the eigenvalues of $\Sigma(F)$, and define

$$\mathbf{D}(F) = \text{diagonal }(d_1, \cdots, d_{p+r}). \tag{5.6}$$

Let

$$\mathbf{G}(F) = \begin{pmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{pmatrix}: (p + r) \times (p + r), \quad \mathbf{G}_{11}{:}p \times p \tag{5.7}$$

satisfy

$$\mathbf{G}^T(F)\mathbf{G}(F) = \mathbf{I}_{p+r} = \mathbf{G}(F)\mathbf{G}^T(F), \quad \Sigma(F) = \mathbf{G}(F)\mathbf{D}(F)\mathbf{G}^T(F). \tag{5.8}$$

Condition (5.6) ensures the decomposition (5.8) is unique.

The parameter $\theta = (\alpha, \mathbf{B})^T$ can be written as a functional of the unknown distribution function F,

$$\theta(F) = ((-\mathbf{B}(F), \mathbf{I}_r)\mu(F), \quad \mathbf{G}_{21}(F)\mathbf{G}_{11}^{-1}(F))^T, \tag{5.9}$$

assuming that $\mathbf{G}_{11}^{-1}(F)$ exists.

Assuming the common distribution of the $\mathbf{e}_i$ to be multivariate normal, Gleser (1981) obtained the m.l.e. of θ in the functional equations model (i.e., where the u_i are fixed not random). This is given by

$$\hat{\theta} = \begin{pmatrix} \hat{\alpha} \\ \hat{\mathbf{B}} \end{pmatrix} = \begin{pmatrix} \alpha(F_n) \\ \mathbf{B}(F_n) \end{pmatrix} = \theta(F_n) = \begin{pmatrix} (-B(F_n), I_r)\mu(F_n) \\ \mathbf{G}_{21}(F_n)\mathbf{G}_{11}^{-1}(F_n) \end{pmatrix}, \tag{5.10}$$

assuming that $\mathbf{G}_{11}^{-1}(F_n)$ exists. Gleser also discusses the uniqueness properties of this estimator and examines its relation to ordinary and generalized least squares estimators using orthogonally invariant norms. It is easily seen that one obtains the same m.l.e. in the structural equations model, with the extra assumption that the common distribution of the $\mathbf{u}_i$ is multivariate normal.

Note that $\theta(F_n)$ can be regarded as the method of moments estimator for all distribution functions F with finite high order moments.

5.2 *Calculation of the influence function.* The methods of perturbation analysis are used to derive the influence function.

As in Section 1, let $W = (1 - \varepsilon)F + \varepsilon\delta_{\mathbf{z}}$. Define $\mathbf{D}(W)$, $\mathbf{G}(W)$, $\Sigma(W)$ as in equations (5.6)–(5.8) and $\theta(W) = (\alpha(W), \mathbf{B}(W))^T$ by (5.9). Then, assuming existence of $G_{11}^{-1}(W)$ for ε small,

$$\begin{aligned} IC(\mathbf{B}, F, z) &= \frac{d}{d\varepsilon}\mathbf{B}(W)\Big|_{\varepsilon=0} = \frac{d}{d\varepsilon}[\mathbf{G}_{21}(W)\mathbf{G}_{11}^{-1}(W)]\Big|_{\varepsilon=0} \\ &= \left(\frac{d}{d\varepsilon}\mathbf{G}_{21}(W)\Big|_{\varepsilon=0} - \mathbf{B}(F)\frac{d}{d\varepsilon}\mathbf{G}_{11}(W)\Big|_{\varepsilon=0}\right)\mathbf{G}_{11}^{-1}(F). \end{aligned} \tag{5.11}$$

To evaluate $(d/d\varepsilon)\mathbf{G}_{21}(W)\,|_{\varepsilon=0}$, expand $\Sigma(W)$, $\mathbf{G}(W)$ and $\mathbf{D}(W)$ as functions of ε and use the identities (5.8) and

$$\Sigma(W) = (1 - \varepsilon)\Sigma(F) + \varepsilon(1 - \varepsilon)(\mathbf{z} - \mu_X(F))(\mathbf{z} - \mu_X(F))^T, \tag{5.12}$$

to obtain the following formula for the (i, j)th element of $\mathbf{G}^{(1)} = d\mathbf{G}(W)/d\varepsilon$ at $\varepsilon = 0$.

(5.13) $$(\mathbf{G}^{(1)})_{ij} = \sum_{k \neq i}([\mathbf{G}^T(F)(\mathbf{z} - \mu_X(F))][\mathbf{G}^T(F)(\mathbf{z} - \mu_X(F))]^T(\mathbf{G})_{kj})/(d_k - d_i).$$

Then $IC(\mathbf{B}, F; z)$ is given by (5.11) and (5.13).

For the intercept we have

(5.14) $$IC(\alpha(F), \mathbf{z}) = (-\mathbf{B}(F), \mathbf{I}_r)(\mathbf{z} - \mu_X(F)) - \mu_{X_1}(F)IC(\mathbf{B}(F), \mathbf{z}).$$

It is not difficult to check that the influence function is translation and scale invariant. Further computational details may be found in Kelly (1981, pages 49–54).

The asymptotic covariance matrix $\Sigma(\theta(F_n))$, as in Section 1, is

(5.15) $$\Sigma(\theta(F_n)) = E_F\{IC(\theta(F_n), \mathbf{z})IC^T(\theta(F_n), \mathbf{z})\}, \\ = \begin{pmatrix} \Sigma(\alpha(F_n)) & \Sigma(\alpha(F_n), \mathbf{B}(F_n)) \\ & \Sigma(\mathbf{B}(F_n)) \end{pmatrix}.$$

For F the multivariate normal the elements of (5.15) can be simplified to

(5.16) $$\Sigma(\alpha(F_n)) = \sigma^2\{1 + \mu_1^T[\sigma^2\, \Sigma_{\mathrm{U}}^{-1}(\mathbf{I}_p + \mathbf{B}^T\mathbf{B})^{-1}\Sigma_{\mathrm{U}}^{-1} + \Sigma_{\mathrm{U}}^{-1}]\mu_1\}(\mathbf{I}_r + \mathbf{B}^T\mathbf{B}).$$

The elements of $\Sigma(\alpha(F_n), \mathbf{B}(F_n))$: $r(p + 1) \times r(p + 1)$ giving the asymptotic covariance between the (i, j)th elements of $\sqrt{n}(\mathbf{B}(F_n) - \mathbf{B})$ and the ℓth element of $\sqrt{n}(\alpha(F_n) - \alpha)$ are equal to

(5.17) $$\sigma^2(\mathbf{I}_r + \mathbf{B}^T\mathbf{B})_{i\ell}\{[\sigma^2\, \Sigma_{\mathrm{U}}^{-1}(\mathbf{I}_p + \mathbf{B}^T\mathbf{B})^{-1}\, \Sigma_{\mathrm{U}}^{-1} + \Sigma_{\mathrm{U}}^{-1}]\mu_1\}_j.$$

The elements of $\Sigma(\mathbf{B}(F_n))$: $rp \times rp$ giving the asymptotic covariance between the (i, j)th and (i', j')th elements of $\sqrt{n}(\mathbf{B}(F_n) - \mathbf{B})$ are

(5.18) $$\sigma^2[\sigma^2\, \Sigma_{\mathrm{U}}^{-1}(\mathbf{I}_p + \mathbf{B}^T\mathbf{B})^{-1}\Sigma_{\mathrm{U}}^{-1} + \Sigma_{\mathrm{U}}^{-1}]_{jj'}[\mathbf{I}_r + \mathbf{B}^T\mathbf{B}]_{ii'}.$$

As in Section 2, nonparametric estimators of $\Sigma(\theta(F_n))$ are $\hat{\Sigma}_1$, $\hat{\Sigma}_J$ where

(5.19) $$\hat{\Sigma}_1 = \sum_{i=1}^n \frac{IC(\theta(F_n), \mathbf{z}_i)IC^T(\theta(F_n), \mathbf{z}_i)}{n}$$

and $\hat{\Sigma}_J$ is given by (3.2).

6. Numerical results.

6.1 *Example.* Miller (1980, pages 127–142) presented simultaneous pairs of measurements of serum kanamycin levels in blood samples drawn from twenty babies. One of the measurements was obtained by a heelstick method (X), the other using an umbilical catheter (Y). The question was whether the catheter value systematically differed from the heelstick value and so could it be substituted for it after correction for bias. Since there was measurement error in both methods, an errors in variables rather than regression analysis was used. It was reasoned that $\lambda = 1$ was correct.

The twenty pairs of heelstick and catheter values are presented in Table 1. The estimates from the analysis were

(6.1) $\hat{\alpha} = -1.16, \quad \hat{\beta} = 1.07, \quad \hat{\sigma}_1^2 = \hat{\sigma}_2^2 = 4.60, \quad \hat{\sigma}_{U_1} = 21.4, \quad \hat{\sigma}_{U_2}^2 = 24.5$

Estimates of the standard error (s.e.) of $\hat{\alpha}$ and $\hat{\beta}$ discussed in Section 2 are presented in Table 2. The bootstrap values were calculated using the following algorithm (cf., Efron, 1981, page 591).

1. Fit the sample distribution function F_{20} mass 1/20 at

(6.2) $$\mathbf{z}_i = (x_i, y_i)^T, \ i = 1, \cdots, 20.$$

2. Draw a "bootstrap sample" from F_{20}, $\mathbf{z}_1^*, \cdots, \mathbf{z}_{20}^*$ i.i.d. F_{20} and calculate

(6.3) $$\hat{\boldsymbol{\theta}}^* = \hat{\theta}(\mathbf{z}_1^*, \cdots, \mathbf{z}_{20}^*) = (\hat{\alpha}^*, \hat{\beta}^*)^T.$$

3. Independently repeat step 2, $B(B = 200$ or $500)$ times obtaining "bootstrap

TABLE 1
Serum kanamycin levels in blood samples drawn simultaneously from an umbilical catheter and a heel venapuncture in twenty babies

Baby	**Heelstick**	**Catheter**
1	23.0	25.2
2	33.2	26.0
3	16.6	16.3
4	26.3	27.2
5	20.0	23.2
6	20.0	18.1
7	20.6	22.2
8	18.9	17.2
9	17.8	18.8
10	20.0	16.4
11	26.4	24.8
12	21.8	26.8
13	14.9	15.4
14	17.4	14.9
15	20.0	18.1
16	13.2	16.3
17	28.4	31.3
18	25.9	31.2
19	18.9	18.0
20	13:8	15.6

TABLE 2
Standard errors of the estimators $\hat{\alpha}$, $\hat{\beta}$.

Estimator	**Influence function: $\hat{\Sigma}_1$**	**Jackknife: $\hat{\Sigma}_J$**	**Normal Theory: $\hat{\Sigma}_1$**	**Bootstrap $B = 200$**	**Bootstrap $B = 500$**
$\hat{\alpha}$	4.06	4.97	3.39	4.58	4.33
$\hat{\beta}$	.21	.26	.16	.23	.22

replications" $\hat{\theta}^{*1}, \cdots, \hat{\theta}^{*B}$, and calculate

$$\text{S.E.}(\hat{\beta}) = [\textstyle\sum_{b=1}^{B} (\hat{\beta}^{*b} - \hat{\beta}^{*\cdot})^2/(B-1)]^{1/2} \tag{6.3}$$

where $\hat{\beta}^{*\cdot} = \sum_{b=1}^{B}(\hat{\beta}^{*b})/100$. The quantity S.E. $(\hat{\alpha})$ is computed similarly. The 95% normal confidence interval for $\beta(F)$ using $\hat{\Sigma}_2$ is (.76, 1.38). Another method for constructing a $100(1-\alpha)\%$ confidence interval for $\beta(F)$ based on normal theory assumptions is outlined in Kendall and Stuart (1961, pages 388–390). The interval is the transform of

$$\begin{aligned} &\tan^{-1}(\beta/\sqrt{\lambda}) \in \tan^{-1}(\beta(F_n)/\lambda) \\ &\qquad \pm \frac{1}{2}\sin^{-1}\left\{2t_{n-2}^{\alpha/2}\left[\frac{(S_X^2S_Y^2 - S_{XY}^2)}{(n-2)(S_Y^2 - \lambda S_X^2)^2 + 4\lambda S_{XY}^2}\right]^{1/2}\right\} \end{aligned} \tag{6.5}$$

where $t_{n-2}^{\alpha/2}$ is the $1-(\alpha/2)$ percentile point of the t-distribution with $(n-2)$ degrees of freedom. In this example (6.5) yields the interval (.76, 1.52). All the 95% normal confidence intervals using the estimates of s.e. of Table 2 support the hypothesis $\alpha = 0.0$, $\beta = 1.0$, i.e., the two different methods of measurement are equivalent. We see that the normal theory and influence function estimates of s.e. are the smallest. In the Monte Carlo experiment of the next section this also occurs.

Detection of influential observations. Table 3 presents the jackknife and sample influence functions of the slope and intercept, at each data point. We see the estimates are very similar. The values of the normalized influence function, J_i and V_i are also tabled. Babies 2 and 16 have the largest values of the influence function and have a negative influence on the slope estimate. When omitted from the analysis the estimate of slope is increased. Removal of Baby 2 moves the estimate of $\boldsymbol{\theta}$ to approximately the edge of a 60% confidence region around $\hat{\boldsymbol{\theta}}$, since $1.536 = \chi_2^2(.4)$. Removal of Baby 16 moves the estimate of $\boldsymbol{\theta}$ to the edge of a 40% confidence region around $\hat{\boldsymbol{\theta}}$.

Monte Carlo estimators of the estimators of standard deviation of $\hat{\alpha}$ and $\hat{\beta}$. To further illustrate the procedures for estimating the standard deviation (s.d.) of $\hat{\alpha}$ and $\hat{\beta}$ presented in Section 2, Monte Carlo studies similar to Efron (1981) were carried out. Each study comprised 200 trials of (x_i, y_i) where $x_i = u_{1i} + e_{1i}$ and $y_i = \alpha + \beta u_{1i} + e_{2i}$, $i = 1, \cdots, 20$. Tables 4–6 show summary statistics over the 200 trials.

The last line of each table gives the "true" standard deviation of $\hat{\alpha}$ and $\hat{\beta}$. This was in fact estimated using the sample standard deviation from a separate Monte Carlo experiment of 2,000 trials, in each situation.

In Table 4 the asymptotic variances given by expressions (2.11) and (2.12) are valid (since the underlying distribution has finite fourth moments) and are calculated. For example, line 3 of Table 4 shows that the normal theory estimates for the 200 trials averaged .0614, with sample standard deviation .0014, and coefficient of variation $.1863 = .0614/.0014$. The value $\sigma(\hat{\alpha})$ is .6625. The root mean squared error, from .0625, was $\sqrt{\text{m.s.e.}} = .0115$. For each table, the bootstrap

TABLE 3
Estimates of the influence function for each data point

Baby	Intercept: $IC_J(\alpha, F_n; \mathbf{z}_i)$ Jackknife Influence	Intercept: $IC(\alpha, F_n; \mathbf{z}_i)$ Sample Influence	Slope: $IC_J(\beta, F_n; \mathbf{z}_i)$ Jackknife Influence	Slope: $IC(\beta, F_n; \mathbf{z}_i)$ Sample Influence	J_i Normalised Jackknife Influence	V_i Normalized Sample Influence
1	−3.73	−3.68	0.26	.26	.032	.024
2	77.9	61.66	−4.27	−3.36	1.336	1.536
3	−1.73	−1.64	0.07	.06	.004	.002
4	−1.15	−1.06	0.07	.08	.004	.0
5	1.08	1.07	0.09	.09	.122	.102
6	−6.35	−6.34	0.20	.20	.106	.092
7	.77	.77	0.03	.03	.028	.024
8	−7.45	−7.35	0.27	.26	.098	.082
9	3.43	3.36	−0.12	−0.12	.020	.020
10	−14.71	−14.71	0.52	0.52	.396	.350
11	8.53	8.02	−0.53	−0.49	.052	.050
12	−10.75	−10.85	0.74	0.74	.222	.176
13	4.61	4.21	−0.19	−0.17	.018	.018
14	−15.54	−14.93	0.62	0.59	.016	.234
15	−6.35	−6.34	0.20	0.20	.106	.092
16	26.85	23.83	−1.11	−0.98	.534	.580
17	−19.38	−16.20	1.05	0.88	.090	.090
18	−33.12	−30.47	1.85	1.68	.248	.290
19	−3.80	−3.75	0.13	0.13	.028	.026
20	16.15	14.40	−0.67	−0.59	.198	.206

TABLE 4.

Method	$\hat{\alpha}$ AVE	SD	CV	$\sqrt{\text{m.s.e.}}$	$\hat{\beta}$ AVE	SD	CV	$\sqrt{\text{m.s.e.}}$
Influence Function	.062	.013	.21	.013	.063	.022	.35	.022
Jackknife	.068	.015	.22	.016	.076*	.030	.39	.032
Normal Theory	.061	.011	.19	.012	.067	.020	.30	.020
Bootstrap, $B = 200$	.065	.015	.23	.016	.074*	.026	.35	.028
Bootstrap, $B = 500$	.065	.013	.21	.014	.074*	.025	.34	.027
Asymptotic Value	.063				.064			
True Value	.063				.065			

standard deviations in each trial were calculated according to the algorithm given by (6.2)–(6.4).

Tables 4 through 6 are a comparison of methods of assigning standard errors to $\hat{\alpha}$ and $\hat{\beta}$. Each table consists of 200 trials of $x_i = u_{1i} + e_{1i}$ and $y_i = \alpha + \beta u_{1i} + -e_{2i}$, $i = 1, \cdots, 20$ with $\alpha = 0.0$ and $\beta = .5$.

In Table 4: $e_{1i} \sim .25N(0, 1)$, $e_{2i} \sim .25N(0, 1)$, $u_{1i} \sim N(0, 1)$.

In Table 5: $e_{1i} \sim .25N(0, 1)$, $e_{2i} \sim .25N(0, 1)$, $u_{1i} \sim$ Cauchy(0, 1).

In Table 6: $e_{1i} \sim .25(\exp(1) - 1.0)$, $e_{2i} \sim .25(\exp(1) - 1.0)$, $u_{1i} \sim$ Cauchy(0, 1).

In Tables 4 through 6 large biases are indicated by asterisks: *Relative bias ≥ 0.10, **Relative bias ≥ 0.20, ***Relative bias ≥ 0.40.

TABLE 5

Method	$\hat{\alpha}$ AVE	SD	CV	$\sqrt{\text{m.s.e.}}$	$\hat{\beta}$ AVE	SD	CV	$\sqrt{\text{m.s.e.}}$
Influence Function	.062	.011	.18	.011	.008***	.009	1.11	.011
Jackknife	.066	.013	.19	.013	.016	.014	.88	.014
Normal Theory	.061	.012	.19	.012	.010**	.009	.85	.010
Bootstrap, $B = 200$	.065	.012	.19	.012	.017**	.010	.58	.011
Bootstrap, $B = 500$	.065	.012	.19	.012	.017**	.010	.58	.011
True Value	.063				.014			

TABLE 6

Method	$\hat{\alpha}$ AVE	SD	CV	m.s.e.	$\hat{\beta}$ AVE	SD	CV	m.s.e.
Influence Function	.060	.016	.25	.017	.007***	.008	1.04	.011
Jackknife	.065	.017	.26	.017	.016	.013	.82	.013
Normal Theory	.061	.024	.39	.025	.010**	.009	.84	.010
Bootstrap, $B = 200$	.062	.017	.27	.017	.017*	.010	.61	.010
Bootstrap, $B = 500$	.063	.017	.26	.017	.017*	.010	.59	.010
True Value	.066				.015			

The obvious conclusion to be drawn from these tables is that the influence function estimate of the standard error is consistently biased downward. Efron (1981) found a similar result in the case of estimation of the variability of a correlation coefficient. In Table 5 it underestimates the s.e. ($\hat{\beta}$) by a factor of two almost and the bias there constitutes 70% of the $\sqrt{\text{m.s.e.}}$ Since the influence function estimate uses sample moments to estimate population moments, bias is to be expected in many sampling situations. The fact that this bias is always downward remains unexplained.

The normal theory estimate performed well in the normal situation but otherwise was biased downward.

The bootstrap performed better than the jackknife in terms of $\sqrt{\text{m.s.e.}}$ However, both were biased upward, the relative bias for the bootstrap in Table 5 being greater than 20%. Examination of some of the "bootstrap replications" of (5.4) revealed that extreme values of these replications occurred. These led to a large value for the average of the bootstrap standard deviations. This leads to the question raised by Efron (1981) as to the purpose of estimating a standard error. If the distribution of $\hat{\beta}$ is not symmetric the standard error is not a good summary measure and it is appropriate to capture any asymmetry in presenting confidence intervals. The best way to construct a confidence interval remains an open problem.

Table 4 and similar studies revealed that the asymptotic standard errors of $\hat{\beta}$ were always less than the "true" values. Perhaps taking higher order terms in the asymptotic calculations might give considerably better approximations. However, this does not explain the bias downward of the influence function estimator.

In summary, it seems that the influence function estimator, i.e., the delta method, should be avoided. The jackknife estimator is conservative in that it overestimates the s.e. on the average but it has a large standard deviation. The bootstrap does better than the jackknife in terms of $\sqrt{\text{m.s.e.}}$ but it can also give biased results.

REFERENCES

ARVESON, J. N. (1969). Jackknifing *U*-statistics. *Ann. Math. Statist.* **40** 2076–2100.

ATKINSON, A. C. (1982). Regression diagnostics, transformations and constructed variables (with discussion). *J. Roy. Statist. Soc. B* **44** 1–36.

BRILLINGER, D. R. (1966). Discussion on Mr. Sprent's paper. *J. Roy. Statist. Soc. B* **28** 294.

CHUNG, K. L. (1974). *A Course in Probability Theory*, Second Edition. Academic, New York.

COOK, R. D. (1977). Detection of influential observations in linear regression. *Technometrics* **19** 15–18.

COOK, R. D. (1979). Influential observations in linear regression. *J. Amer. Statist. Assoc.* **74** 169–174.

COOK, R. D. and WEISBERG, S. (1982). *Residuals and Influence in Regression.* Chapman and Hall, New York.

DEVLIN, S. J., GNANADESIKAN, R., and KETTENRING, J. R. (1975). Robust estimation and outlier detection with correlation coefficients. *Biometrika* **62** 531–545.

EFRON, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7** 1–26.

EFRON, B. (1980). The jackknife, the bootstrap, and other resampling plans. Technical Report No. 163, Department of Statistics, Stanford University.

EFRON, B. (1981). Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. *Biometrika* **68** 589–599.

EFRON, B. and STEIN, C. (1978). The jackknife estimate of variance. Technical Report No. 40, Division of Biostatistics, Stanford University.

GLESER, L. J. (1981). Estimation in a multivariate "errors in variables" regression model: Large sample results. *Ann. Statist.* **9** 24–44.

HANDSCHIN, E., KOHLAS, J., FIECHTER, A. and SCHWEPPE, F. (1975). Bad data analysis of power system state estimation. *IEEE Transactions on Power Apparatus and Systems*, PAS-94, **2** 329–337.

HINKLEY, D. V. (1977). On jackknifing in unbalanced situations. *Technometrics* **19** 285–292.

HINKLEY, D. V. (1978). Improving the jackknife with special reference to correlation estimation. *Biometrika* **65** 13–21.

HUBER, P. J. (1977). *Robust Statistical Procedures.* SIAM, Philadelphia.

HUBER, P. J. (1980). Lecture notes on robust regression and covariance. Summer Research Conference and Numerical and Statistical Analysis, June 15–27, 1980, University of Delaware.

KELLY, G. E. (1981). The influence function in the errors in variables problem. Technical Report No. 68, Division of Biostatistics, Stanford University.

KENDALL, M. G. and STUART, A. (1961). *The Advanced Theory of Statistics*, Vol. 2. Hafner, New York.

KRASKER, W. S. and WELSCH, R. E. (1979). Efficient bounded-influence estimation using alternative definitions of sensitivity. Technical Report No. 3, Alfred P. Sloan School of Management, Massachusetts Institute of Technology, Cambridge.

MADANSKY, A. (1959). The fitting of straight lines when both variables are subject to error. *J. Amer. Statist. Assoc.* **54** 173–205.

MILLER, R. G., JR. (1974). The jackknife—a review. *Biometrika* **61** 1–15.

MILLER, R. G., JR. (1980). Kanamycin levels in premature babies. Biostatistics Casebook, Vol. III, 127–142. Technical Report No. 57, Division of Biostatistics, Stanford University.

REIERSØL, O. (1950). Identifiability of a linear relation between variables which are subject to error. *Econometrica* **18** 375–389.

DEPARTMENT OF STATISTICS
UNIVERSITY COLLEGE
CORK, IRELAND