MEASUREMENT ERROR MODEL COLLINEARITIES

by

Olivia Carrillo-Gamboa and Richard F. Gunst

Technical Report No. SMU/DS/TR/240
Department of Statistical Science

June 1990

Measurement Error Model Collinearities

Olivia Carrillo-Gamboa and Richard F. Gunst

Department of Statistical Science
Southern Methodist University
Dallas, TX 75275-0332

ABSTRACT

Collinearities for linear and nonlinear measurement error models are defined as small norms of second-order moment matrices. For linear measurement error models, collinearities are a property of the second-order moment matrix of the unobservable true predictor variables. For nonlinear measurement error models, collinearities are a property of the second-order moment matrix of derivatives of the nonlinear function. Diagnostics for the detection of collinearities are presented. These diagnostics are similar to those used in traditional regression models, only they are applied to estimated second-order moment matrices. Examples are discussed, one of which shows that measurement errors can mask collinearities among the true predictors.

KEY WORDS: Diagnostics; Errors in variables; Regression.

1. INTRODUCTION

Classical assumptions for regression models (e.g., Gunst and Mason 1980, Section 6.1) require that the predictor variables be nonstochastic and measured without error. When these assumptions are reasonable, least squares estimation is applied to estimate model parameters and to draw inferences. While one can still use least squares estimation and invoke conditional inferences when the predictors are stochastic, conditional inferences do not obviate the need for alternatives to least squares when predictors are measured with error (e.g., Fuller 1987, Section 1.1). Measurement error modeling is, therefore, replacing least squares estimation in many applications in which predictor variables are contaminated by measurement error (e.g., Battese, Harter, and Fuller 1988; Burr 1988; Carroll and Spiegelman 1986; Carroll, Spiegelman, Lan, Bailey, and Abbott 1984; Gunst and Kelly 1990; Hung and Fuller 1987; Hwang 1986; Stefanski and Carroll 1985).

Recent theoretical work in measurement error modeling is rapidly expanding its application. Nonlinear models (e.g., Fuller 1987, Chapter 3; Wolter and Fuller, 1982), generalized linear models (e.g., Schafer 1987, Stefanski 1989; Stefanski and Carroll 1987), robust regression (e.g., Amman and

Van Ness 1989, Brown 1982; Zamar 1989), and quasilikelihood estimation (e.g., Whittemore and Keller 1988) are but a few of the current research areas being investigated.

This paper focuses on a serious problem for traditional regression models that also affects measurement error modeling procedures: collinearities among the error-free predictors. Collinearities can occur with any linear or nonlinear measurement error models fit by maximum likelihood, generalized likelihood, or quasilikelihood methods. Collinearities are fundamentally a property of the predictors; none of these estimation methods necessarily protect against the occurrence of collinearities or ameliorate their effects.

Collinearities in linear measurement error models are defined in Section 2. Estimators of model parameters are presented in Section 3. In Section 4, procedures for detecting collinearities in linear measurement error models are discussed. The results presented extend to maximum likelihood estimation for generalized linear models and to quasilikelihood estimation, with suitable modifications to account for the weight matrices used in each. Section 5 expands the discussion of collinearities to nonlinear measurement error models. Examples are presented in Section 6 and concluding remarks are made in Section 7.

2. DEFINING COLLINEARITIES

Traditional definitions of collinearities (e.g., Gunst 1983) must be modified for measurement error models in order to accommodate the errors in the true variates. In this section, such modifications are presented for ultrastructural linear measurement error models.

Define a linear measurement error model in terms of unobservable response ψ and predictor $\pi' = (\pi_1, \pi_2, \ldots, \pi_k)$ variates as $\psi = \pi'\beta + q$, where q denotes equation error. If $\sigma_{qq} > 0$, the model is referred to as an equation-error model; whereas if $\sigma_{qq} = 0$ (i.e., $q = 0$), it is referred to as a no-equation-error model (Fuller 1987, Chapter 2). The observable model is $y = x'\beta + e$, where $y = \psi + v$, $x = \pi + u$, with v and u denoting measurement errors. The model error term is $e = q + v - u'\beta$. Let $z = (y\ x')'$, $\xi = (\psi\ \pi')'$, and $w = (v\ u')'$. Then $z = \xi + (q\ 0')' + w$.

For ease of presentation, both the functional (fixed π) and the structural (stochastic π) models are represented by the following ultrastructural model assumption for a sample of size n (see Dolby 1976):

$$\pi_i \sim \mathrm{ID}(\mu_i, \Sigma_{\pi\pi}) \quad i = 1,\ldots,n\ . \tag{2.1}$$

To accommodate the functional model, the degenerate normal distribution with $\Sigma_{\pi\pi} = \Phi$ is permitted. The structural model occurs when $\mu_i = \mu_\pi$ and $\Sigma_{\pi\pi} > \Phi$. In all models, it is assumed that (all limits are taken as $n \rightarrow \infty$)

$$\Gamma_{\pi\pi} = \lim n^{-1} \sum \pi_i \pi_i' = \lim \Gamma_{nn} \tag{2.2}$$

exists, is at least positive semidefinite, and has strictly positive diagonal elements. Derivations of asymptotic distributional properties of measurement error model estimators ordinarily require that the second-order moment matrix $\Gamma_{\pi\pi}$ be positive definite; however, the definition and detection of collinearities only requires that it be positive semidefinite.

The error vectors are assumed to follow a common distribution, independently of the π_i:

$$w_i \sim \mathrm{IID}(0, \Sigma_{ww}) \quad i = 1, \ldots, n \quad , \tag{2.3}$$

where

$$\Sigma_{ww} = \begin{pmatrix} \sigma_{vv} & \Sigma_{vu} \\ \Sigma_{uv} & \Sigma_{uu} \end{pmatrix} \quad ,$$

and, independent of the π_i and the w_i,

$$q_i \sim \mathrm{IID}(0, \sigma_{qq}) \quad .$$

Where needed for moment properties, the variates π_i in (2.1) and the errors w_i and q_i are assumed to follow normal distributions. The normality assumption can be relaxed for the derivation of many asymptotic properties (see Fuller 1987, Chapter 2).

Collinearities are defined in terms of the unobservable predictors. Unobserved rather than observed variates are used in the definition of measurement error model collinearities because measurement errors may mask collinearities. The concrete compressive strength example in Section 6 illustrates such masking.

Definition 2.1

Let $\pi_1, \pi_2, \ldots, \pi_k$ denote predictors following the ultrastructural model assumptions (2.1). Let $P_{\pi\pi}$ be the scaled second-order moment matrix,

$$P_{\pi\pi} = D_{\pi\pi}^{-1/2} \Gamma_{\pi\pi} D_{\pi\pi}^{-1/2} \quad , \tag{2.4}$$

where $\Gamma_{\pi\pi}$ is defined by (2.2) and $D_{\pi\pi}$ is a diagonal matrix containing the diagonal elements of $\Gamma_{\pi\pi}$. If for a suitably small $\eta > 0$, there exists a k-dimensional vector $c \neq 0$ such that

$$P_{\pi\pi}\, c = \delta_\pi \quad \text{with} \quad || \, \delta_\pi \, || < \eta \, |c| \, | \quad , \tag{2.5}$$

then a collinearity exists among the unobservable predictors.

Unlike the definitions of collinearities in traditional regression settings, this definition is posed as an asymptotic result rather than for a fixed sample size n. It could be stated as a finite-sample result by defining a scaled second-order sample moment matrix P_{nn} using Γ_{nn} in (2.2). The definition using (2.5) permits the specification of diagnostic procedures that have reasonable asymptotic properties. It also establishes the uniqueness of collinearities through the use of a fixed vector c, rather than one that might depend on the sample size.

The following lemma provides a necessary and sufficient condition for the existence of a measurement error model collinearity. The proof of this lemma and other theoretical results in this and subsequent sections is contained in the Appendix.

Lemma 2.1

A necessary and sufficient condition for a collinearity among the unobservable predictors satisfying the ultrastructural model assumptions (2.1) is that $\lambda_{min} < \eta$, where λ_{min} is the smallest eigenvalue of $P_{\pi\pi}$.

In many applications of regression modeling in the physical and engineering sciences, interest is in whether collinearities occur among the nonconstant predictor variables. In such cases it is common to investigate the correlation matrix of the nonconstant predictors. Inequality (2.5) is necessarily satisfied when collinearities occur among the nonconstant predictors, as the following theorem states.

Theorem 2.1

Let $R_{\pi\pi}$ denote the correlation matrix among k-1 nonconstant predictors satisfying the ultrastructural model assumptions (2.1). If for a suitably small $\eta > 0$ there exists a k-1 dimensional vector $d \neq 0$ such that

$$R_{\pi\pi}\, d = \alpha_{\pi} \quad \text{with} \quad ||\, \alpha_{\pi} \,|| < \eta \,||\, d \,|| \;, \tag{2.6}$$

then there exists a k-dimensional vector $c \neq 0$ such that (2.5) is satisfied.

3. ESTIMATORS

In order to accommodate both functional and structural models through the ultrastructural framework of (2.1), moment estimators will be used to estimate the model parameters. Define $M_{xx} = n^{-1} \sum x_i x_i'$, and similarly for the matrices M_{uu}, $M_{u\pi}$, and $M_{\pi u}$. Assume that the error covariance matrix, Σ_{ww} in (2.3), is either known or unbiasedly estimated by a sample covariance matrix which follows a multiple of a Wishart distribution having d_f degrees of freedom. In the estimator formulas that follow, let S_{ww} denote either the known or estimated covariance matrix. In the no-equation error

formulation of the model, Σ_{ww} need not be completely known. If $\Sigma_{ww} = \sigma_{ww}T_{ww}$, only T_{ww} need be known.

An estimator of the second-order moment matrix is $\tilde{\Gamma}_{\pi\pi} = M_{xx} - S_{uu}$ for equation-error models and $\tilde{\Gamma}_{\pi\pi} = M_{xx} - \tilde{\lambda}S_{uu}$, where $\tilde{\lambda}$ is the smallest root of $| M_{zz} - \lambda S_{ww}| = 0$, for no-equation-error models. If Σ_{ww} is only known up to a multiple for no-equation-error models, T_{ww} replaces S_{ww} in the estimator and in the determination of $\tilde{\lambda}$. Let $\tilde{D}_{\pi\pi} = \mathrm{diag}(\tilde{\Gamma}_{\pi\pi})$. Then an estimator of the scaled second-order moment matrix is:

$$\tilde{P}_{\pi\pi} = \tilde{D}_{\pi\pi}^{-1/2}\, \tilde{\Gamma}_{\pi\pi}\, \tilde{D}_{\pi\pi}^{-1/2} \quad . \tag{3.1}$$

Application of Chebychev's theorem and of stochastic order properties for sample moments of (not necessarily iid) random variables establishes the following properties of the estimators:

$$\tilde{\Gamma}_{\pi\pi} = \Gamma_{nn} + O_p(n^{-1/2}) \tag{3.2}$$

$$\tilde{P}_{\pi\pi} = P_{nn} + O_p(n^{-1/2}) \quad .$$

Assuming equation error, the vector of regression coefficients can be estimated as (e.g., Fuller 1987, Section 2.2.1):

$$\tilde{\beta} = (M_{xx} - S_{uu})^{-1}(M_{xy} - S_{uv}) \quad . \tag{3.3}$$

Under the no-equation error model assumptions, the estimator has a similar form (e.g., Fuller 1987, Section 2.3.1):

$$\tilde{\beta} = (M_{xx} - \tilde{\lambda}S_{uu})^{-1}(M_{xy} - \tilde{\lambda}S_{uv}) \quad , \tag{3.3}$$

where $\tilde{\lambda}$ is the smallest root of $| M_{zz} - \lambda S_{ww} | = 0$. If Σ_{ww} is only known up to a multiple, the estimator of β has the same form as (3.3) with the elements of Tww replacing the corresponding ones of Sww and with $\tilde{\lambda}$ the smallest root of $| M_{zz} - \lambda T_{ww}| = 0$. All of these estimators can be written in a common form:

$$\tilde{\beta} = \tilde{\Gamma}_{\pi\pi}^{-1}\, \tilde{\Gamma}_{\pi\psi} \quad . \tag{3.4}$$

Under very general assumptions, $n^{1/2}(\tilde{\beta} - \beta)$ is asymptotically normally distributed. Assuming normally distributed errors, the covariance matrix of this asymptotic distribution is

$$\Omega_{\beta\beta} = \Gamma_{\pi\pi}^{-1} \{(\sigma_{qq} + \sigma_{rr})(\Gamma_{\pi\pi} + \Sigma_{uu})$$

$$+ c(1 + \nu)\, \Sigma_{ue}\, \Sigma_{eu} + \nu\sigma_{rr}\, \Sigma_{uu}\}\, \Gamma_{\pi\pi}^{-1}\,, \tag{3.5}$$

where $\sigma_{rr} = \sigma_{vv} - 2\Sigma_{vu}\beta + \beta'\Sigma_{uu}\beta$, c = 1 for equation-error models, and c = -1 and $\sigma_{qq} = 0$ for no-equation-error models. The scalar $\nu = \lim\,(n/d_f)$ if Σ_{ww} is estimated and $\nu = 0$ if Σ_{ww} is known or known up to a multiple. This covariance matrix can be consistently estimated (Fuller 1987, Chapter 2). Of particular interest to this investigation is the insertion of $\tilde{\Gamma}_{\pi\pi}$ as an estimator of $\Gamma_{\pi\pi}$ in equation (3.5).

The deleterious effects of collinearities on measurement error model estimators and asymptotic properties are similar to their well-known effects on least squares estimation (e.g., Gunst 1983). Briefly, the asymptotic variance formula (3.5) involves the inverse of the estimated second-order moment matrix, $\tilde{\Gamma}_{\pi\pi}$. The existence of collinearities among the unobservable true predictors will cause inflation of the variances. Expressing the estimators in the common form (3.4) shows that signs and magnitudes of coefficient estimates can be influenced by the presence of collinearities through the inverse of the estimated second-order moment matrix. These characteristics and many others parallel those for least squares. They are clearly illustrated in the examples in Section 6.

4. COLLINEARITY DETECTION

Rather than develop entirely new diagnostics for collinearity detection in measurement error models, the purpose of this section is to establish that diagnostics that are already available for least squares estimators can be applied to measurement error model estimation. One reason for preferring such an approach is that if measurement errors are sufficiently small, the diagnostics will reduce to their appropriate least squares equivalents. A second reason is that standard algorithms, perhaps with minor modification, can be used to calculate the diagnostics.

Virtually all of the collinearity diagnostics that have been developed for least squares estimators can be applied in the measurement error model setting to the estimated second-order moment matrix or to one of its scaled or standardized modifications. The justification for this statement is the following theorem, which can be proven using consistency arguments based on equations (3.2).

Theorem 4.1

Let $\{\pi_i\}$ be a sequence of k-dimensional predictors following the ultrastructural model assumptions of Section 2. Let $P_{\pi\pi}$ be the scaled second-order moment matrix, defined by equation (2.4), and $\tilde{P}_{\pi\pi}$ its consistent estimator, defined by equation (3.1). If a collinearity exists among the predictors, with c defined as in (2.5), then

$$\tilde{P}_{\pi\pi}c = \delta_n \quad \text{with} \quad \text{plim}|\,|\,\delta_n\,|\,| < \eta\,|\,|\,c\,|\,| \quad . \tag{4.1}$$

Theorem 4.1 establishes that the criterion (2.5) will be satisfied with probability arbitrarily close to 1 for a sufficiently large sample size. Alternatively, one can establish that the difference between $P_{\pi\pi}c = \delta_\pi$ and $\tilde{P}_{\pi\pi}\ c = \delta_n$ is a term of $O_p(n^{-1/2})$. The proof of this result differs for functional and structural model assumptions, but both proofs are based on limit properties of second-order moments.

Theorem 4.2

Under either the functional or the structural model assumptions, if a collinearity exists among the predictors, with c defined as in (2.5), then

$$\tilde{P}_{\pi\pi}c = \delta_n \quad \text{with} \quad |\,|\,\delta_n\,|\,| < \eta\,|\,|\,c\,|\,| + O_p(n^{-1/2}) \quad . \tag{4.2}$$

Parallel results hold for the sample correlation matrix of the nonconstant predictor variables:

$$\tilde{R}_{\pi\pi} = \tilde{D}_{11}^{-1/2}(\tilde{S}_{xx} - S_{11})\tilde{D}_{11}^{-1/2} \quad , \tag{4.3}$$

where

$$\tilde{D}_{11} = \text{diag}(\tilde{S}_{xx} - S_{11})$$

$$\tilde{S}_{xx} = n^{-1}\ \Sigma(x_{i1} - \bar{x}_1)\ (x_{i1} - \bar{x}_1)'$$

and $\bar{x}_1 = n^{-1}\sum x_{i1}$, and x_{i1} is the ith vector of nonconstant predictor variables; i.e., $x_i' = (1\ x_{i1}')$. The matrix S_{11} is the lower (k-1) x (k-1) portion of the known (Σ_{ww}) or estimated (S_{ww}) error covariance matrix.

In the examples in Section 6, least squares collinearity diagnostics are applied to the sample correlation matrix (4.3). These diagnostics and others not illustrated in the examples can be applied to the scaled second-order moment matrix if collinearities involving the constant term are to be assessed. The importance of Theorems 4.1 and 4.2 is that they establish the ability to detect collinearities among the true unobservable predictors π_i from sample second-order moment matrices based on sufficiently large sample sizes.

5. NONLINEAR MEM MODELS

The results presented in the previous sections are readily extended to a variety of alternative models. In this section, specific application is made to implicit nonlinear measurement error models of

the form $f(\xi,\beta) = 0$, where $\xi = (\psi\ \pi')'$ is again the vector of unobservable response and predictor variables.

Britt and Luecke (1973) (see also Fuller 1987, Section 3.2) present an iterative estimation scheme for the no-equation-error model with predictors satisfying the functional model assumptions and normally distributed errors. Let $f(\xi_t, \beta_t)$, $f_\beta(\xi_t, \beta_t)$, and $f_\xi(\xi_t, \beta_t)$ denote, respectively, the nonlinear implicit function and column vectors of its first derivatives with respect to the unknown coefficients and the unknown variates. The updated estimator for the regression coefficients can then be expressed as:

$$\begin{aligned}\tilde{\beta}_{t+1} = \tilde{\beta}_t &+ \{\sum f_\beta(\tilde{\xi}_{ti}, \tilde{\beta}_t)\ \tilde{\sigma}_{tii}^{-1} f_\beta(\tilde{\xi}_{ti}, \tilde{\beta}_t)'\}^{-1} \\ &\cdot \{\sum[\tilde{e}_i + f(\tilde{\xi}_{ti}, \tilde{\beta}_t)]\ \tilde{\sigma}_{tii}^{-1} f_\beta(\tilde{\xi}_{ti}, \tilde{\beta}_t)\}\ ,\end{aligned} \tag{5.1}$$

where $\tilde{e}_i = f_\xi\ (\tilde{\xi}_{ti}, \tilde{\beta}_t)'(z_i - \tilde{\xi}_{ti})$ and $\tilde{\sigma}_{tii} = f_\xi(\tilde{\xi}_{ti}, \tilde{\beta}_t)'\Sigma_{ww} f_\xi(\tilde{\xi}_{ti}, \tilde{\beta}_t)$.

In equation (5.1), the derivatives $f_\xi(\tilde{\xi}_{ti}, \tilde{\beta}_t)$ perform the role that estimated variate values do in linear measurement error model estimation. The estimated second-order moment matrix $\tilde{\Gamma}_{\pi\pi}$ in equation (3.4) is thereby replaced by the weighted sum of the products of the derivatives in the second term of (5.1). This motivates the following definition for collinearities in implicit nonlinear measurement error models.

Definition 5.1

Assume an implicit nonlinear functional relationship of the form $f(\xi,\beta) = 0$. Let $f_\beta(\xi,\beta)$ denote the first partial derivatives of $f(\xi,\beta)$, evaluated at the true parameter values and $\sigma_{ii} = f_\xi(\xi_i, \beta)'$ $\Sigma_{ww} f_\xi(\xi_i,\beta)$. Assume that

$$\Gamma_{\pi\pi} = \lim n^{-1} \sum f_\beta(\xi_i,\beta)\ \sigma_{ii}^{-1} f_\beta(\xi_i,\beta)' \tag{5.2}$$

exists, is at least positive semidefinite, and has strictly positive diagonal elements. A collinearity exists among the variates in the nonlinear model if for a suitably small $\eta > 0$ there exists a nonzero vector of constants c such that

$$P_{\pi\pi}\ c = \delta_\pi \quad \text{with} \quad ||\ \delta_\pi\ || < \eta\ ||\ c\ ||\ , \tag{5.3}$$

where $P_{\pi\pi} = D_{\pi\pi}^{-1/2}\ \Gamma_{\pi\pi}\ D_{\pi\pi}^{-1/2}$ and $D_{\pi\pi} = \text{diag}(\Gamma_{\pi\pi})$.

This definition does not define a collinearity among the predictor variables as does Definition 2.1. The existence of a collinearity is model dependent, through the partial derivatives, and may not

indicate any relationship among the predictors themselves. A parallel in linear regression is the occurrence of collinearities among polynomial functions of predictors in polynomial regression when no strong collinearities exist among the predictors themselves (e.g., Bradley and Srivastava 1979). Linear measurement error model collinearity diagnostics can be adapted to nonlinear models. Convergence and consistency properties of the estimators (5.1) are assumed in this context, properties which can be proven as in Fuller (1987, Section 3.2), which also contains estimating equations for the true variate values ξ_i. The following theorem generalizes the results of Section 4 to implicit nonlinear measurement error models.

Theorem 5.1

Let $(\tilde{\xi}_{ti}, \tilde{\beta}_t)$ denote the estimates after t iterations. Define, for a fixed sample size n, the estimated weighted second-order moment matrix of partials from the tth iteration

$$\tilde{\Gamma}_{n,t} = n^{-1} \sum f_\beta(\tilde{\xi}_{ti}, \tilde{\beta}_t)\, \tilde{\sigma}_{tii}^{-1} f_\beta(\tilde{\xi}_{ti}, \tilde{\beta}_t)' , \tag{5.4}$$

and assume that the diagonals of $\tilde{\Gamma}_{n,t}$ are strictly positive for all (n,t). Let $\tilde{\Gamma}_{n,\infty}$ denote the estimated second-order moment matrix using the converged estimates $(\tilde{\xi}_i, \tilde{\beta})$. Assume that $\Gamma_{\pi\pi} = \lim \tilde{\Gamma}_{n,\infty}$. If a collinearity exists in an implicit nonlinear measurement error model, then for n and t sufficiently large

$$\tilde{P}_{n,t} c = \delta_{n,t} \quad \text{with} \quad \text{plim} \,||\, \delta_{n,t} \,|| < \eta \,||\, c \,|| \quad , \tag{5.3}$$

where $\tilde{P}_{n,t} = \tilde{D}_{n,t}^{-1/2}\, \tilde{\Gamma}_{n,t} \tilde{D}_{n,t}^{-1/2}$ and $\tilde{D}_{n,t} = \text{diag}(\tilde{\Gamma}_{n,t})$.

Amemiya and Fuller (1988) propose a second-order bias adjustment to the nonlinear estimator. If this adjustment is made, the principal results of this section remain valid, although some definitional and implementation details must be modified. Diagnostics computed on the matrix of first partial derivatives can still be used to screen for collinearities.

6. EXAMPLES

6.1 Concrete Compressive Strength Data

Figure 1 is a plot of compressive strength (psi) measurements of samples of concrete two and seven days after pouring. The data were collected in order to investigate prediction equations for

strength measurements of the samples twenty-eight days after pouring. One sample in the original data set is clearly an outlier and has been removed; 40 observations remain and are displayed in Table 1.

The plotted points do not suggest that the two predictors are collinear. Least squares collinearity diagnostics reinforce this supposition: the pairwise correlation between the variates is .72, the smallest eigenvalue of the correlation matrix of the predictors is .28, and the variance inflation factors are 2.10. The least squares estimates and their estimated standard errors are shown in Table 2. Calculated t statistics for the predictors indicate that the Day 2 strength measurements can be deleted from the model (p = .24).

A structural no-equation-error measurement error model can also be fit to the data with the assumption that the measurement errors are uncorrelated and the error variances are equal; i.e., $\Sigma_{WW} = \sigma_{WW}I$ with σ_{WW} unknown. The estimated regression coefficients are shown in Table 2 along with their estimated standard errors. The estimates are strikingly different from the least squares estimates. In particular, the coefficient estimates are an order of magnitude larger, approximately equal, and opposite in sign. The estimated standard errors are almost two orders of magnitude larger than those for least squares. These are the classical symptoms of collinearities.

Using the correlation matrix of the estimated error-free true predictors, the collinearity diagnostics clearly indicate severe collinearity between them: the pairwise correlation between the variates is .994, the smallest eigenvalue of the correlation matrix of the predictors is .006, and the variance inflation factors equal 85. Figure 2 is a plot of the estimated error-free predictor values. The plotted points fall very near to a straight line, visually confirming the presence of the collinearity.

6.2 Prosthetic Hip Data

An example of an implicit nonlinear measurement error model is the prosthetic hip data of Reilly and Patino-Leal (1981) (see also Fuller 1987, p.245). The data are x-y coordinates of an x-ray image of implanted hip prostheses. A five-parameter ellipse of the form

$$\beta_3(y - \beta_1)^2 + 2\beta_4(y - \beta_1)(x - \beta_2) + \beta_5(x - \beta_2)^2 - 1 = 0$$

was fit to the data using the methodology of Section 5. The parameter estimates and their estimated standard errors are shown in Table 3. These estimates are based on the fitting methodology of Section 3. Bias-adjusted estimates were also computed. They lead to substantively the same conclusions as the unadjusted ones.

The variance inflation factors in Table 3 suggest the presence of a collinearity among the at least the first two partial derivatives. This is not surprising because both derivatives are linear combinations of $y-\beta_1$ and $x-\beta_2$. While there is no extremely large correlation between pairs of the estimated derivatives (maximum correlation = .72), other diagnostics confirm the presence of the collinearity. For example. the smallest eigenvalue of the correlation matrix of the derivatives is .018. Its corresponding

eigenvector is (.59 -.58 .32 -.34 .33). Thus, all the partial derivatives are indeed collinear with one another.

7. CONCLUDING REMARKS

The presence of collinearities among the true error-free predictors poses interesting questions for the conduct of measurement error modeling. Accommodation of collinearities can be accomplished in many ways, from the deletion of variables to the use of traditional biased regression estimators such as ridge regression (e.g., Jagpal 1982). The assumptions traditionally made with measurement error models imply that any collinearity would be population-inherent (e.g., Gunst 1983). If so, variable selection is a viable accommodation strategy.

One must, however, recognize the burden placed on the traditional assumptions when small data sets are fit with measurement error model techniques. The two examples presented in Section 6, in spite of their small sample sizes, can be reasonably argued to contain population-inherent collinearities, the first because of the measurement process used and the second because of the polynomial model used. In the first example, deletion of one of the predictors is defensible. In the second, there is no need to delete variates if the primary purpose is to fit the ellipse and not to provide a meaningful interpretation of the individual coefficients. The question of estimation alternatives for collinearities that are not population-inherent merits theoretical investigation.

ACKNOWLEDGEMENTS

The authors express their appreciation to Richard M. Weed, New Jersey Department of Transportation, for permission to use the concrete compressive strength data.

APPENDIX: PROOFS

Proof of Lemma 2.1:

A collinearity exists if (2.5) holds. Then

$$|| \delta_\pi \ ||^2 = \delta_\pi{}'\delta_\pi = c'P^2_{\pi\pi}c.$$

By the Courant-Fischer min-max theorem (Fuller 1987, p. 391):

$$\lambda_{\min}(P^2_{\pi\pi}) \, || \, c \, ||^2 \leq c'P^2_{\pi\pi}c = || \, \delta_\pi \, ||^2 < \eta^2 || \, c \, ||^2.$$

Necessity follows because $P_{\pi\pi}$ is symmetric and at least positive semidefinite, implying that $\lambda_{\min}(P^2_{\pi\pi}) = \{\lambda_{\min}(P_{\pi\pi})\}^2$. Sufficiency follows by letting c be the eigenvector corresponding to

$\lambda_{min}(P_{\pi\pi})$.

The proof of Theorem 2.1 uses the following lemma.

Lemma A

Let $G = D^{1/2}E^{-1/2} = \mathrm{diag}(g_{11}, \ldots, g_{mm})$ and $H = D^{-1/2}E^{1/2} = \mathrm{diag}(h_{11}, \ldots, h_{mm})$, with $g_{jj} \geq 1$ and $0 < h_{jj} \leq 1$. Let v be any m-dimensional real vector. Then (i) $||Gv|| \geq ||v||$ and (ii) $||Hv|| \leq ||v||$.

Proof of Lemma A:

$$||Gv||^2 = \sum |g_{jj}v_j|^2 \geq \sum |v_j|^2 = ||v||^2. \text{ Similarly for (ii).}$$

Proof of Theorem 2.1:

Let the first element of π represent the constant term, 1. Partition $\Gamma_{\pi\pi}$ as follows:

$$\Gamma_{\pi\pi} = \begin{pmatrix} 1 & \mu_1' \\ \mu_1 & \Gamma_{11} \end{pmatrix},$$

where $\mu_\pi = (1\ \mu_1') = \lim n^{-1}\sum \pi_i$. Let $D_{\delta\delta}$ contain the k-1 elements of $D_{\pi\pi}$ correponding to the nonconstant predictors. Define $\Delta_{\pi\pi} = \Gamma_{\pi\pi} - \mu_\pi\mu_\pi'$, with Δ_{11} denoting the lower (k-1) x (k-1) corner of $\Delta_{\pi\pi}$. Let D_{11} denote a diagonal matrix containing the diagonal elements of Δ_{11}. Then

$$P_{\pi\pi} = \begin{pmatrix} 1 & \mu_1'D_{\delta\delta}^{-1/2} \\ D_{\delta\delta}^{-1/2}\mu_1 & D_{\delta\delta}^{-1/2}\Gamma_{11}D_{\delta\delta}^{-1/2} \end{pmatrix}$$

and $R_{\pi\pi} = D_{11}^{-1/2}\,\Delta_{11}D_{11}^{-1/2} = D_{11}^{-1/2}(\Gamma_{11} - \mu_1\mu_1')\,D_{11}^{-1/2}$.

Next, let $c' = (c_0\ c_1')$, where $c_1 = D_{\delta\delta}^{1/2}D_{11}^{-1/2}d$ and $c_0 = -\mu_1'D_{\delta\delta}^{-1/2}c_1 = -p_1'c_1$. Then

$$\delta = \begin{pmatrix} 0 \\ (-p_1p_1' + P_{11})c_1 \end{pmatrix},$$

where P_{11} is the lower k×k corner of $P_{\pi\pi}$. Now, $||c|| \geq ||c_1||$ and from Lemma A(i), $||c_1|| \geq ||d||$. Straightforward algebra yields

$$(-p_1p_1' + P_{11})c_1 = D_{\delta\delta}^{-1/2}D_{11}^{-1/2}R_{\pi\pi}d \quad ,$$

so that, invoking Lemma A(ii), $||\delta_\pi|| \leq ||\alpha_\pi||$. The conclusion follows from the premise of the theorem since $||c|| \geq ||d||$.

Proof of Theorem 5.1:

Let $|| \delta_\pi || = (\eta - \eta_0) || c ||$ for some $0 < \eta_0 < \eta$. Since $\Gamma_{n,\infty}$ converges to $\Gamma_{\pi\pi}$ and $P_{n,t}$ is a continuous function of the elements of $\Gamma_{n,t}$, for every $\delta > 0$ there exist n_δ, and t_n such that for $n > n_\delta$ and $t \geq t_n$, $\Pr\{ || (P_{n,t} - P_{\pi\pi}) c || < \eta_0 || c || \geq 1-\delta$. Consequently, with probability arbitrarily close to one, $|| \delta_{n,t} || = || P_{n,t}c || \leq || (P_{n,t} - P_{\pi\pi})c || + || P_{\pi\pi}c || < \eta_0 || c || + (\eta - \eta_0) || c || = \eta || c ||$.

REFERENCES

Amemiya, Y. and Fuller, W.A. (1988). "Estimation for the Nonlinear Functional Relationship," *The Annals of Statistics*, 16, 147-160.

Ammann, L. and Van Ness, J.V. (1989). "Standard and Robust Orthogonal Regression," *Communications in Statistics -Simulation and Computation*, 18, 145-162.

Battese, G.E., Harter, R.M., and Fuller, W.A. (1988). "An Error-Components Model for Prediction of County Crop Areas Using Survey and Satellite Data," *Journal of the American Stastistical Association*, 83, 28-36.

Bradley, R.A. and Srivastava, S.S. (1979). "Correlation in Polynomial Regression," *The American Statistician*, 33, 11-14.

Britt, H.I. and Luecke, R.H. (1973). "The Estimation of Parameters in Nonlinear Implicit Models," *Technometrics*, 15, 233-247.

Brown, M. L. (1982). "Robust Line Estimation with Errors in Both Variables," *Journal of the American Statistical Association*, 77, 71-79; Corrigenda, 78, 1008.

Burr, D. (1988), "On Errors-in-Variables in Binary Regression -- Berkson Case," *Journal of the American Statistical Association*, 83, 739-743.

Carroll, R.J. and Spiegelman, C.H. (1986). "The Effect of Ignoring Small Measurement Errors in Precision Instrument Calibration," *Journal of Quality Technology*, 18, 170-173.

Carroll, R.J., Spiegelman, C.H., Lan, K.K., Bailey, K.T., and Abbott, R.D. (1984). "On Errors-in-Variables for Binary Regression Models," *Biometrika*, 71, 19-25.

Dolby, G.R. (1976). "The Ultra-Structural Relation: A Synthesis of the Functional and Structural Relations," *Biometrika*, 63, 39-50.

Fuller, W.A. (1987). Measurement Error Models, New York: John Wiley and Sons, Inc.

Gunst, R.F. (1983). "Regression Analysis with Multicollinear Predictor Variables: Definition, Detection, and Effects," *Communications in Statistics -- Theory and Methods*, 12, 2217-2260.

Gunst, R.F. and Kelly, N.A. (1990). "Captive-Air Irradiation Experiments on Ozone Formation in Southern California," Research Publication GMR 7055, General Motors Research Laboratories, Warren, MI.

Gunst, R.F. and Mason, R.L. (1980). *Regression Analysis and Its Application*, New York: Marcel Dekker, Inc.

Hung, H-M and Fuller, W.A. (1987). "Regression Estimation of Crop Acreages with Transformed Landsat Data as Auxiliary Variables," *Journal of Business & Economic Statistics*, 5, 475-482.

Hwang, J.T. (1986), "Multiplicative Errors-in-Variables Models with Applications to Recent Data Released by the U.S. Department of Energy," *Journal of the American Statistical Association*, 81, 680-688.

Jagpal, H.S. (1982). "Multicollinearity in Structural Equation Models with Unobservable Variables," *Journal of Marketing Research*, 19, 431-439.

Reilly, P.M. and Patino-Leal, H. (1981). "A Bayesian Study of the Error-in-Variable Model," *Technometrics*, 23, 221-231.

Schafer, D.W. (1987). "Covariate Measurement Error in Generalized Linear Models," *Biometrika*, 74, 385-391.

Stefanski, L.A. (1989), "Correcting Data for Measurement Error in Generalized Linear Models," *Communications in Statistics - Theory and Methods*, 18, 1715-1733.

Stefanski, L.A. and Carroll, R.J. (1985), "Covariate Measurement Error in Logistic Regression," *The Annals of Statistics*, 13, 1335-1351.

Stefanski, L.A. and Carroll, R.J. (1987), "Conditional Scores and Optimal Scores for Generalized Linear Measurement-Error Models," *Biometrika*, 74, 703-716.

Whittemore, A.S. and Keller, J.B. (1988). "Approximations for Regression with Covariate Measurement Error," *Journal of theAmerican Statistical Association*, 83, 1057-1066.

Wolter, K.M. and Fuller, W.A. (1982). "Estimation of Nonlinear Errors-in-Variables Models," *Annals of Statistics*, 10, 539-548.

Zamar, R.H. (1989). "Robust Estimation in the Errors-in-Variables Model," *Biometrika*, 76, 149-160.

Fig. 1 Scatterplot of Day 2 and Day 7 Strength Measurements.

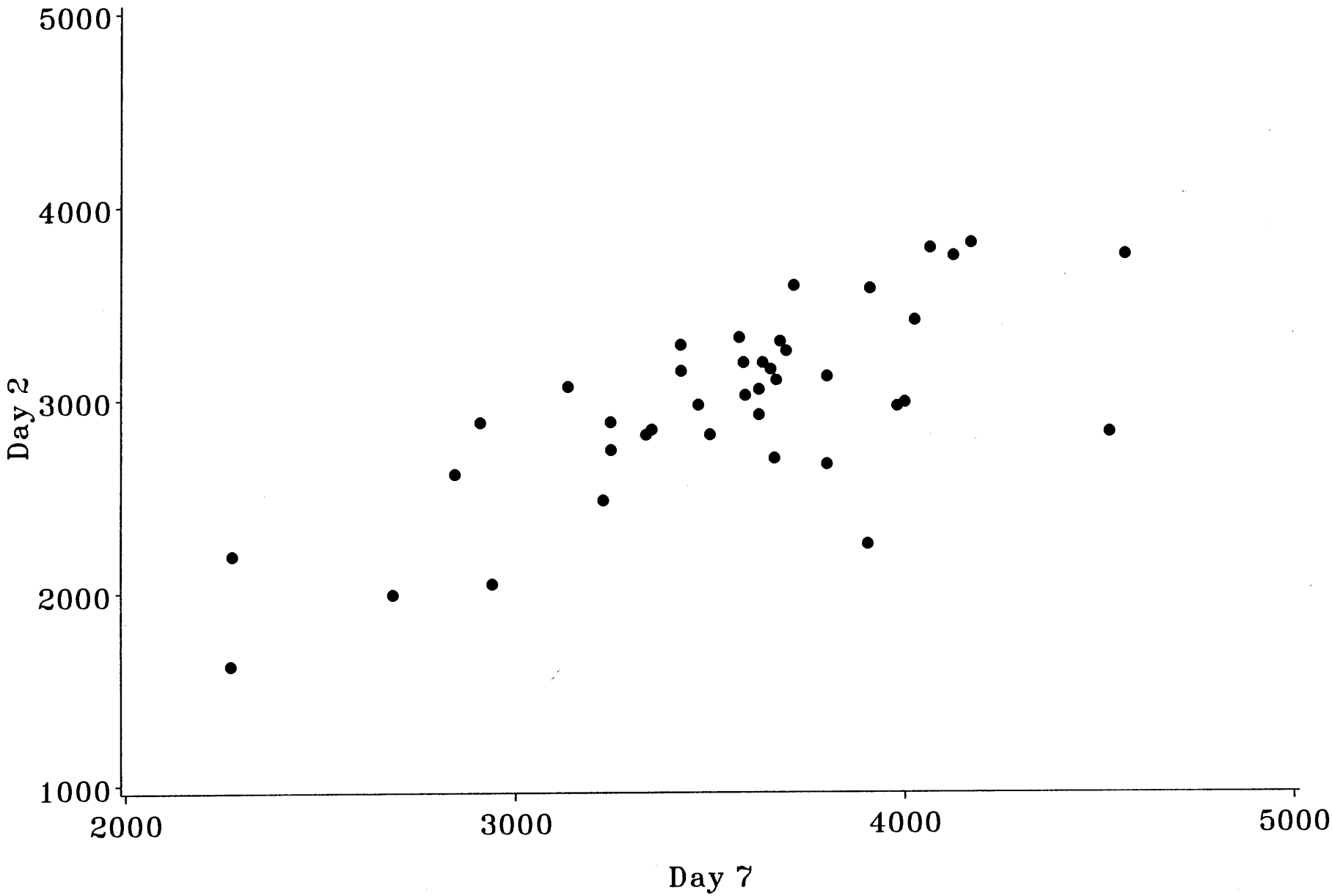

Fig. 2 Scatterplot of Estimated Day 2 and Day 7 Strength Measurements.