Sankhyā: The Indian Journal of Statistics
1993, Special Volume 55, Series B, Pt. 3, pp. 353–371
Dedicated to the memory of P. C. Mahalanobis

FINITE SAMPLE PROPERTIES OF A MEASURE OF MULTICOLLINEARITY[1]

By J. S. MEHTA
Temple University, U.S.A.
P. A. V. B. SWAMY
Federal Reserve Board, U.S.A.
and
N. S. IYENGAR
Indian Statistical Institute, Bangalore

SUMMARY. This article deals with exact finite sample properties of a summary statistic that is used to measure the effects of multicollinearity on the contributions of regressors to the square of the multiple correlation coefficient of a regressand with the regressors. This statistic measures the "distance" of a cross-product matrix from the diagonal matrix obtained by zeroing its off-diagonal elements and hence is useful in detecting near multicollinearities in regression problems. It can also distinguish between apparent and real multicollinearities with positive probability.

1. Introduction

It is often desired to analyze the effects of multicollinearities on regression results. In an important paper, Stewart (1987) has carefully compared several measures of multicollinearity and showed that the square roots of variance inflation factors, unlike the other measures he considered, are invariant under column scaling and can measure the distance between a cross-product matrix and a matrix that is exactly singular. One would like to measure not only this distance but also the distance between a cross-product matrix and the diagonal matrix obtained by zeroing its off-diagonal elements. Therefore, in this paper we present a formula that can measure the latter distance and study its properties.

In Section 2, a measure of multicollinearity effect considered previously by Theil (1971) is stated in different but equivalent forms. One of these forms shows that this measure can indicate in a precisely quantifiable manner how near a cross-product matrix is to the diagonal matrix obtained by zeroing its off-diagonal elements. After deriving this measure's first four moments, a study of its exact finite sample distribution in the normal case is also carried out. Lastly, some concluding remarks are presented in Section 3.

[1]Views expressed in this article are those of the authors and do not necessarily reflect those of the Board of Governors or the staff of the Federal Reserve System, U.S.A.

2. A STOCHASTIC MULTICOLLINEARITY MEASURE AND ITS PROPERTIES

2.1 *The model.* Let a linear regression model be given by

$$y = X\beta + u, \tag{1}$$

where y is a T-vector of observations on a dependent variable, X is a $T \times K$ matrix of observations on K independent variables, β is a K-vector of regression coefficients, and u is a T-vector of disturbances. (Note : With an abuse of notation we use the same symbol to denote a random variable and the value assumed by it.) It is assumed that u is a normal variable with $\mathcal{E}(u|X) = 0$, $\mathcal{E}(u\,u'|X) = \sigma^2 I$, and $rank\,(X) = K$.

2.2 *Theil's measure of multicollinearity and its generalizations.* To know the effects of the nondiagonality of $X'X$ on the contributions of the columns of X to the square of the multiple correlation coefficient of y with X, Theil (1971, p. 179) considers a multicollinearity measure which is

$$\tilde{m} = R^2 - \sum_{j=1}^{k}(R^2 - R^2_{-j}), \tag{2}$$

where R^2 is the square of the multiple correlation coefficient of y with the columns of X, and R^2_{-j} is the square of the multiple correlation coefficient of y with the columns of X_{-j} which is X without X's jth column. The quantity $(R^2 - R^2_{-j}) \geq 0$ is termed the "incremental contribution" of the jth regressor by Theil. If we use a Bayes estimator of β in place of the least squares estimator of β used in (2), then we get Swamy, Mehta, Thurman and Iyengar's (1985) generalized multicollinearity index. This index may be of interest to those who believe that Bayes estimators produce very stable parameter estimates.

2.3 *Analytical properties of* $\tilde{m}$. Defining $R^2 = b'X'y/y'y$ and

$R^2_{-j} = y'X_{-j}(X'_{-j}X_{-j})^{-1}X'_{-j}y/y'y$, we can write

$$\tilde{m} = \frac{y'Qy}{y'y}, \tag{3}$$

where $Q = X(X'X)^{-1}X' - \sum_{j=1}^{K}[X(X'X)^{-1}X' - X_{-j}(X'_{-j}X_{-j})^{-1}X'_{-j}]$ which, in view of the matrix identity given in Theil (1971, p. 682, (B.23)), simplifies to $Q = X(X'X)^{-1}X' - \sum_{j=1}^{K} M_{-j}x_j(x_j'M_{-j}x_j)^{-1}x_j'M_{-j}$ if $M_{-j} = I - X_{-j}(X'_{-j}X_{-j})^{-1}X'_{-j}$. The vector $M_{-j}x_j$ may be recognized as the residual vector resulting from the least squares fit of the jth column of X on its remaining columns. Letting $e_j = M_{-j}x_j$, we obtain $Q = X(X'X)^{-1}X' - \sum_{j=1}^{K} e_j(e_j'e_j)^{-1}e_j'$.

In their discussion of Stewart's (1987, p. 95) paper, Hadi and Velleman point out that plots of the ith diagonal element of $e_j(e_j'e_j)^{-1}e_j'$ versus the jth diagonal element of $X_{-j}(X'_{-j}X_{-j})^{-1}X'_{-j}$. for each j, can detect multicollinearities and multicollinearity-influential points. Since the matrix Q compares each element of $X(X'X)^{-1}X'$ with the corresponding element of $\sum_{j=1}^{k} e_j(e'_j e_j)^{-1}e'_j$, it combines the merits of Hadi and Velleman's scatterplots in a single diagnostic measure.

Many feel that when there is a constant term in model (1), model (1) should be centred before the measure (2) is applied. So a restatement of $\tilde{m}$ in terms of centred data may be useful. Suppose that the first element of β represents the constant term of model (1). This means that a vector of unit elements, denoted by ι, is the first column of X. Let the remaining columns of X be represented by the $T \times (K-1)$ matrix Z. Let Z_{-j} be Z without Z's jth column and let $A = I - (1/T)\iota\iota'$. Then suppressing the contribution of the constant term, we can define R^2 as $y'AZ(Z'AZ)^{-1}Z'Ay/y'Ay$ and R^2_{-j} as $y'AZ_{-j}(Z'_{-j}AZ_{-j})^{-1}Z'_{-j}Ay/y'Ay$. The jth t-ratio, denoted by t_j, may be defined as the ratio of the jth element of $(Z'AZ)^{-1}Z'Ay$ to the square root of the jth diagonal element of $[y'My/(T-K)](Z'AZ)^{-1}$. A useful identity that holds among these variables is $(1-R^2)t_j^2/(T-K) = R^2 - R^2_{-j}$ (see Theil 1971, p. 175). Because of this identity $\tilde{m}$ can be expressed in terms of R^2, R^2_{-j} and t_j as

$$\tilde{m} = \frac{y'AZ(Z'AZ)^{-1}Z'Ay}{y'Ay} - \frac{(1-R^2)}{(T-K)}\sum_{j=1}^{k-1} t_j^2 . \tag{4}$$

The use of the same symbol $\tilde{m}$ to denote the statistics (2) and (4) should not obscure their differences. The former estimates $\mathcal{E}(\tilde{m}|X, \beta, \sigma^2)$ while the latter estimates $\mathcal{E}(\tilde{m}|AZ$, slope coefficients, $\sigma^2)$. It is possible that these two conditional expectations are not equal. The equivalence of the conditions, (i) $X'X$ is diagonal, (ii) $Q = 0$, and (iii) $\tilde{m}$ in (3) is degenerate at 0, is established by Swamy et al. (1985). Equation (4) gives one more condition under which $\tilde{m}$ is degenerate at 0. The usual F and t tests of hypotheses about β are equivalent if $\frac{(T-K)R^2}{(K-1)(1-R^2)} = \frac{1}{(K-1)}\sum_{j=1}^{K-1} t_j^2$. Under precisely this condition, $\tilde{m} = 0$ with probability 1, as can be seen from equation (4). In real econometric situations, one can test whether multicollinearity is a "problem" or not by comparing the conclusions of the F and t tests. For example, if a test based on the F statistic $\frac{(T-K)R^2}{(K-1)(1-R^2)}$ rejects (or accepts) the hypothesis that $\beta_2 = \cdots = \beta_K = 0$ and a test based on the t_j statistic accepts (or rejects) the hypothesis that $\beta_j = 0$, then there is a contradiction. The greater the probability with which $\tilde{m}$ takes large absolute values, the greater the chance of this contradiction occurring. For this reason, the econometrician cannot turn away from $\tilde{m}$ altogether.

An alternative form of equation (4) is

$$\tilde{m} = \frac{(t'D^{\frac{1}{2}}D^{*\frac{1}{2}}CD^{*\frac{1}{2}}D^{\frac{1}{2}}t - t't)y'My}{y'Ay(T-K)}, \tag{5}$$

where D and D^* are the diagonal matrices obtained by zeroing the off-diagonal elements of $(Z'AZ)^{-1}$ and $(Z'AZ)$, respectively, $C = D^{*-\frac{1}{2}}Z'AZD^{*-\frac{1}{2}}$ is a correlation matrix, $(1 - R^2) = y'My/y'Ay$, and

$$t = (t_1, t_2, \cdots, t_{K-1})' = \sqrt{\frac{T-K}{y'My}} D^{-\frac{1}{2}}(Z'AZ)^{-1}Z'Ay\,.$$

The standard decomposition, $y'Ay = y'AZ(Z'AZ)^{-1}Z'Ay + y'My$, shows that the factor $y'My/[y'Ay(T-K)]$ in (5) is positive and less that 1 with probability 1. The first term on the right hand side of this decomposition divided by $y'My/(T-K)$ can be expressed as a quadratic form in t, as shown in (5). Substituting these results into equation (5) gives

$$\tilde{m} = \frac{(t'D^{\frac{1}{2}}Z'AZD^{\frac{1}{2}} - I)t}{t'D^{\frac{1}{2}}Z'AZD^{\frac{1}{2}}t + T - K}, \tag{6}$$

which is in the form of a ratio of two quadratic forms in Student's t variates.

1. It now follows form equation (6) that $\tilde{m}$ measures the distance between $D^{\frac{1}{2}}Z'AZD^{\frac{1}{2}}$ and I or between $Z'AZ$ and D^{-1} which is the same as D^* whenever $Z'AZ$ is diagonal. The numerator of the ratio (6) measures the distance between $D^{\frac{1}{2}}Z'AZD^{\frac{1}{2}}$ and I and its denominator has only the effect of making this measure less than one. This means that there is a direct relation between $\tilde{m}$ and near multicollinearity represented by the departures of $Z'AZ$ from D^*.

2. The diagonal elements of $D^{\frac{1}{2}}Z'AZD^{\frac{1}{2}}$ are the same as those of DD^* which are the variance inflation factors (VIF's) computed from centered data (see Stewart 1987, p. 72, (4.1)). The VIF's are greater than or equal to 1 because every diagonal element of D is greater than or equal to the reciprocal of the corresponding diagonal element of D^* (see Rao 1973, p. 74, Problem 20.2(a)).

3. The reason why the VIF's do not measure the distance between the matrices $D^{\frac{1}{2}}Z'AZD^{\frac{1}{2}}$ and I is that DD^* is not equal to $D^{\frac{1}{2}}Z'AZD^{\frac{1}{2}}$ (which is equal to $D^{\frac{1}{2}}D^{*\frac{1}{2}}CD^{*\frac{1}{2}}D^{\frac{1}{2}}$) unless $C = I$. When there is near multicollinearity, $C \neq I$.

4. Even though multicollinearity is present in model (1) whenever $X'X$ is nondiagonal, it is only apparent but not real if the population values of the elements of β corresponding to the nonorthogonal columns of X are all zero. Real multicollinearity arises if $X'X$ is nondiagonal and the population values of at least two elements of β corresponding to the nonorthogonal columns of X are not zero. From this it follows that the VIF's which are functionally independent of β cannot detect real multicollinearity. Since

the nondiagonality of $X'X$ cannot be taken as an indication of the presence of real multicollinearity in the absence of information about β, the dependence of $\tilde{m}$ on both the t-ratios and $Z'AZ$ shown by equation (5) is desirable. The variable $\tilde{m}$ being a function of both the matrix $(D^{\frac{1}{2}}Z'AZD^{\frac{1}{2}} - I)$ and the t-ratios is more comprehensive and more closely related to near multicollinearity than the VIF's if by near multicollinearity economists mean the departures of $D^{\frac{1}{2}}Z'AZD^{\frac{1}{2}}$ from I. We use the values of $\tilde{m}$ to distill a large amount of information about the matrix, $(D^{\frac{1}{2}}Z'AZD^{\frac{1}{2}} - I)$, and the vector β, into a single number. A nonzero value of $\tilde{m}$ can point to real multicollinearity if the F and t_j tests reject the hypotheses that $\beta_2 = \cdots = \beta_K = 0$ and $\beta_j = 0$, respectively.

5. Stewart (1987) likes the VIF's because they are invariant under a scaling of any column of X. The variable $\tilde{m}$ is scale invariant, not only with respect to the scaling of the columns of X but also the scaling of y, as equation (2) shows.

6. While equation (3) proves that $\tilde{m}$ is always bounded between m_L and m_U which are the smallest and the largest eigenvalues of Q, respectively, equation (6) proves that $m_U < 1$. This result was noted previously by Swamy et. al. (1985, p. 411).

7. It can be shown that the trace of Q is zero. This means that m_U and m_L cannot have the same sign. Since m_U is positive, m_L has to be negative. Consequently, $\tilde{m}$ can take negative values with positive probability.

8. Equation (6) implies that $pr(\tilde{m} >, =, \text{ or } < 0) = 1$ according as the matrix $(D^{\frac{1}{2}}Z'AZD^{\frac{1}{2}} - I)$ is positive definite, 0, or negative definite regardless of the value of β. It cannot be shown that real multicollinearity does not arise when the matrix $(D^{\frac{1}{2}}Z'AZD^{\frac{1}{2}} - I)$ is negative definite. Therefore, the fact that $\tilde{m}$ can take negative values with positive probability is of no concern. Since it is the distance between $D^{\frac{1}{2}}Z'AZD^{\frac{1}{2}}$ and I that gives the correct measure of real multicollinearity when the population values of the coefficients corresponding to the nonorthogonal columns of X are nonzero, a value of $\tilde{m}$ does not halve the wrong sign if it is negative.

9. However, we shall see in section 2.7 below that the probability that $m_L \leq \tilde{m} \leq -1$ can be very small even when m_L is much smaller than -1. It is also possible that for $-1 < m_0 < m_1 < 1, pr(m_0 \leq \tilde{m} \leq m_1) \approx 1$ (see Sections 2.4 and 2.7 below).

10. Replacing the matrix $D^{\frac{1}{2}}Z'AZD^{\frac{1}{2}}$ by the matrix $D^{\frac{1}{2}}D^{*\frac{1}{2}}CD^{*\frac{1}{2}}D^{\frac{1}{2}}$, we can write explicitly the individual terms in the numerator and the denominator of the ratio (6) as

$$\begin{aligned}\tilde{m} &= \left[\sum_{i=1}^{K-1}(d_i d^* - 1)t_i^2 + \sum_{i=1}^{K-1}\sum_{j(\neq i)=1}^{K-1} r_{ij}t_it_j\sqrt{d_id_i^*d_jd_j^*}\right] \\ &\div \left[\sum_{i=1}^{K-1} d_id_i^*t_i^2 + \sum_{i=1}^{K-1}\sum_{j(\neq i)=1}^{K-1} r_{ij}t_it_j\sqrt{d_id_i^*d_jd_j^*} + (T-K)\right],\end{aligned}$$

where d_i and d_i^* are the ith diagonal elements of D and D^*, respectively, t_i is the ith element of t, and r_{ij} is the ijth element of C. Using the identity $d_id_i^* = (1-R_i^2)^{-1}$, where R_i^2 is the coefficient of determination resulting from the least squares fit of the ith column of X on its remaining columns, established by Stewart (1987, p. 72) and then multiplying both the numerator and the denominator by $\prod_{h=1}^{K-1}(1-R_h^2)$, the ratio $\tilde{m}$ can be rewritten as

$$\begin{aligned}\tilde{m} &= \left[\sum_{i=1}^{K-1}\left(\frac{R_i^2}{1-R_i^2}\right)t_i^2\prod_{h=1}^{K-1}(1-R_h^2) + \sum_{i=1}^{K-1}\sum_{j(\neq i)=1}^{K-1} r_{ij}t_it_j\prod_{h=1}^{K-1}(1-R_h^2)\right. \\ &\times \left.\sqrt{\frac{1}{1-R_i^2}}\sqrt{\frac{1}{1-R_j^2}}\right] \div \left[\sum_{i=1}^{K-1}\frac{t_i^2}{1-R_i^2}\prod_{h=1}^{K-1}(1-R_h^2)\right. \\ &+ \sum_{i=1}^{K-1}\sum_{j(\neq i)=1}^{K-1} r_{ij}t_it_j\prod_{h=1}^{K-1}(1-R_h^2)\sqrt{\frac{1}{1-R_i^2}}\sqrt{\frac{1}{1-R_i^2}} \\ &+ \left.[T-rank(X)]\prod_{h=1}^{K-1}(1-R_h^2)\right],\end{aligned}$$

where K in the definition of t is replaced by $rank(X)$. Now let the hth column of X be linearly dependent on the other columns of X. Then $rank(X) = K-1$, $R_h^2 = 1$, and the hth VIF is equal to ∞. When these results are true, if the t_i's are kept finite and nonzero by using the Moore-Penrose generalized inverse of $Z'AZ$ in place of its regular inverse used in (6), then the terms in the numerator and the denominator of the above ratio become $t_h^2\prod_{j(\neq h)=1}^{K-1}(1-R_j^2)$ if $i = h$ and zero otherwise and $\tilde{m}$ will be equal to 1 with probability 1. This shows that $\tilde{m}$ tends to 1 with probability 1 as near multicollinearity approaches extreme multicollinearity, provided the t-ratios remain finite and nonzero.

Now we numerically evaluate the distribution of $\tilde{m}$ by considering both particular and general cases.

2.4 *Analysis of a particular case of equation* (5). For $K = 3$, equation (5) specializes to

$$\tilde{m} = \frac{[t_1^2(d_1d_1^* - 1) + t_2^2(d_2d_2^* - 1) + 2r\sqrt{d_1d_1^*d_2d_2^*}t_1t_2]y'My}{(T-3)y'Ay}, \tag{7}$$

where d_1 and d_2 are the diagonal elements of D, d_1^* and d_2^* are the diagonal elements of D^*, and r is the correlation coefficient between the two nonconstant regressors. The following results are easy to establish: $(d_1d_1^* - 1) = (d_2d_2^* - 1) = r^2/(1-r^2)$, $\sqrt{d_1d_1^*d_2d_2^*} = (1-r^2)^{-1}$ and $y'Ay = (t_1^2 + 2rt_1t_2 + t_2^2)(1-r^2)^{-1}y'My(T-3)^{-1} + y'My$. Using these results in (7), we find that

$$\tilde{m} = \frac{r(rt_1^2 + 2t_1t_2 + rt_2^2)}{t_1^2 + 2rt_1t_2 + t_2^2 + (T-3)(1-r^2)}. \tag{8}$$

This result makes it clear that when there are two non-constant regressors in model (1), $\tilde{m}$ is degenerate at 0 if $r = 0$ and nondegenerate if $r \neq 0$ regardless of the values of β. When $r = 0$, there is no real multicollinearity. Even though $|r|$ cannot be exactly equal to 1 under our assumption about the rank of X, it may be noted that $pr(\lim_{|r|\to 1}\tilde{m} = 1) = 1$, provided t_1 and t_2 stay finite and nonzero. Both the signs and magnitudes of r, β_2 and β_3 influence the distribution of $\tilde{m}$. Note that the distribution of $\tilde{m}$ depends on β_2 and β_3 only through the means of t_1 and t_2.

To study the distribution of (8) using Monte Carlo methods, we consider the following example : Draper and Smith (1966, p. 366) present the data which they have used to estimate a linear regression model with five nonconstant regressors. Taking two at a time out of these five regressors, we formed ten linear regressions: For $i = 1, 2, \cdots, 10$; $t = 1, 2, \cdots, 13$

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \beta_2 x_{2it} + u_{it}, \tag{9}$$

where the number of observations on the variables of this equation is 13. The correlation coefficients (r_i) between every pair of the nonconstant regressors in these ten regressions are: –.973, –.824, –.821, –.535, –.245, –.139, .030, .229, .731, .816. For each regression in (9), we considered 19 different combinations of the values of β_1 and β_2. β_0 is zero for each of these combinations. One of these combinations was formed so that β_1 and β_2 would be close to zero. The remaining 18 combinations were formed by fixing the value of β_1 (or β_2) at 1 and varying the values of β_2 (or β_1) between –10 and 10. These values are displayed in Table 1.

For each $i = 1, 2, \cdots, 10$, thirteen values of the variables u_{it} were drawn independently from a normal distribution with mean zero and constant variance

Table 1
The intervals (m_0, m_1) for which $pr(m_0 \leq \tilde{m} \leq m_1) = 1$
$\sigma^2 = .2$

β_1	β_2	r: -.973	-.535	.030	.731	.831
1	−10	(.9, 1)	(.3, .35)	(−.05, 0)	(.5, .6)	(.5, .7)
1	−5	(.9, 1)	(.3, .4)	(−.05, 0)	(.45, .5)	(.45, .6)
1	−.2	(.9,1)	(.45, .6)	(−.05, 0)	(−.3, .05)	(.45, .6)
1	.02	(.9,1)	(.15, .35)	(−.05, .05)	(.45, .7)	(.6, .7)
1	.1	(.9,1)	(−.05, .20)	(0, .05)	(.6, .7)	(.6, .8)
1	.5	(.6,.8)	(−.6, −.4)	(0, .05)	(.7, .8)	(.7, .9)
1	1	(−.9, −.6)	(−.2, −.1)	(0, .05)	(.6, .7)	(.8, .9)
1	5	(.9,.1)	(.2, .25)	(0, .05)	(.5, .6)	(.7, .8)
1	10	(.9,.1)	(.2, .3)	(0, .05)	(.5, .6)	(.7, .8)
−10	1	(.9,.1)	(.4, .45)	(−.05, 0)	(.25, .3)	(.6, .7)
−5	1	(.9,.1)	(.45, .5)	(−.05, 0)	(−.2, −.1)	(.5, .6)
−.2	1	(.9, .1)	(.3, .4)	(−.05, 0)	(.4, .6)	(.45 .6)
−.01	1	(.9, .1)	(.25, .35)	(−.05, .05)	(.45, .6)	(.6, .7)
.01	1	(.9, .1)	(.25, .35)	(−.05, .05)	(.5, .6)	(.6, .7)
.1	1	(.9, .1)	(.2, .3)	(0, .05)	(.5, .6)	(.6, .8)
.5	1	(.7, .8)	(0, .15)	(0, .05)	(.6, .7)	(.7, .9)
5	1	(.9, .1)	(−.25, −.15)	(0, .05)	(.7, .8)	(.7, .8)
10	1	(.9, .1)	(.05, .1)	(0, .05)	(.6, .7)	(.7, .8)
.01	.02	(−.7, .8)	(−.5, .45)	(−.05, .05)	(−.5, .7)	(−.3, .8)

$\sigma^2 = 0.2$. For each combination of the values of β_1 and β_2, 130 values of y_{it} were generated using equation (9). This procedure was repeated 3,000 times, giving, for each replication and for each regression, a vector y_i of order 13×1, and a matrix X_i of order 13×3 with a column of ones. This matrix consists of observations on 3 regressors of equation (9). The vector y_i was different for different replications and also for different regressions. But the matrix X_i and the coefficients β_1 and β_2 were different for different regressions and not for different replications.

The 3,000 estimates of $\tilde{m}$ were used to form an empirical conditional distribution of $\tilde{m}$ given $r = r_i, \beta_j$, and $\sigma^2 = 0.2$ for $i = 1, 2, \cdots, 10$; $j = 1, 2, \cdots, 19$, where the number of different combinations of the values of β_1 and β_2 is 19. The intervals (m_0, m_1) for which $pr(m_0 \leq \tilde{m} \leq m_1) = 1$ are presented in Table 1. We have presented these intervals only for 5 values of r: a value near each of $+1$ and -1, one value between -1 and 0, one value between 0 and 1, and one value near 0. It should be noted that since our empirical distributions are only approximations to the exact distribution of $\tilde{m}$, $pr(m_0 \leq \tilde{m} \leq m_1)$ is not exactly equal to 1 and m_0 and m_1, unlike m_L and m_U, are dependent of β. Also, the probabilities with which $\tilde{m}$ takes values in the intervals (m_L, m_0) and (m_1, m_U) may not be exactly zero. The following conclusions emerge from the values given in Table 1:

1. Both m_0 and m_1 are negative in 5 out of 95 cases. Of these 5, 4 correspond to negative r and positive β_1 and β_2 and 1 corresponds to positive r, negative β_1, and positive β_2; m_0 is negative and m_1 is nonnegative in 16 out of 95 cases. In the remaining 74 cases, m_0 is nonnegative and m_1 is positive. This means that $\tilde{m} > 0$ with probability close to 1 in majority of the cases we considered. The probability with which $\tilde{m}$ takes negative values depends on the signs and magnitudes of r, β_1 and β_2.

2. If r is sufficiently away form 0, then the distribution of $\tilde{m}$ is rather sharp, being concentrated about a value that is approximately equal to the magnitude of r, unless the values of both β_1 and β_2 are sufficiently close to 0, in which case the distribution is spread out over an interval around zero. If r is close to zero, then the distribution is close to a degenerate distribution at 0 regardless of the values of β_1 and β_2.

In sum $\tilde{m}$ is a good estimator of real multicollinearity except when $|r| > 0$ and either β_1 or β_2 or both are equal to 0. In these cases $\tilde{m}$ may take nonnegligible values even though real multicollinearity is not present.

2.5 *The exact distribution function of* $\tilde{m}$. The exact distribution function of $\tilde{m}$ is

$$pr(\tilde{m} \leq c) = pr[y'(Q - cI)y \leq 0]$$

$$= pr(\sum_{i=1}^{T} \lambda_i y_i^{*2} \le 0) \; (m_L \le c \le m_U), \tag{10}$$

where $y^* = (y_1^*, y_2^*, \cdots, y_T^*)' = ((1/\sigma)P'y$, P is an orthogonal matrix that diagonalizes the matrix $\sigma^2(Q-cI), \lambda_1 \le \lambda_2 \le \cdots, \lambda_T$ are the eigenvalues of $\sigma^2(Q-cI)$ and $y_1^*, y_2^*, \cdots, y_T^*$ are independent normal deviates, each having unit variance. The mean of y_i^* is μ_i which is the ith element of $(1/\sigma)P'X\beta$. Therefore, y_i^{*2} is a noncentral chi-square variate with 1 degree of freedom and the noncentrality parameter μ_i^2 and the y_i^{*2} are independent. The transformation of $\tilde{m}$ which gives a positive random variable is

$$\hat{m} = \frac{\tilde{m} - m_L}{m_U - m_L}. \tag{11}$$

The exact distribution function of $\tilde{m}$ can be deduced from that of $\hat{m}$ as follows:

$$\begin{aligned} pr(\tilde{m} \le c) &= pr\left(\frac{\tilde{m} - m_L}{m_U - m_L} \le \frac{c - m_L}{m_U - m_L}\right) \\ &= pr(\hat{m} \le d), \end{aligned} \tag{12}$$

where $d = \dfrac{c - m_L}{m_U - m_L}$ and $0 \le d \le 1$.

Following Imhof (1961), we can show that

$$pr(\tilde{m} \le c) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{sin\{\theta(t)\}}{t\rho(t)}dt, \tag{13}$$

where

$$\theta(t) = \frac{1}{2}\sum_{i=1}^{T} \arctan\,(\lambda_i t) + \frac{1}{2}\sum_{i=1}^{T} \frac{\mu_i^2 \lambda_i t}{1+\lambda_i^2 t^2},$$

$$\rho(t) = \prod_{i=1}^{T} (1+\lambda_i^2 t^2)^{\frac{1}{4}} \exp\left\{\frac{1}{2}\sum_{i=1}^{T} \frac{\mu_i^2 \lambda_i^2 t^2}{1+\lambda_i^2 t^2}\right\},$$

and

$$\lim_{t\to 0} \frac{\sin\theta(t)}{t\rho(t)} = \frac{1}{2}\sum_{i=1}^{T} \lambda_i(1+\mu_i^2).$$

2.6 *The first four moments of* $\tilde{m}$. It follows from Mehta and Swamy (1978, Lemma 3, pp. 6–7) that if the rth moment of $\tilde{m}$ exists and if the joint moment generating function, $\Phi(t_1, t_2) = \mathcal{E}\ [\exp(t_1 y'y + t_2 y'Qy)]$, exists and is uniformly continuous for $-\infty < t_1 \leq \epsilon > 0$ and $|t_2| \leq \epsilon > 0$, then

$$\mathcal{E}(\tilde{m}^r) = \frac{1}{\Gamma(r)} \int_{-\infty}^{0} (-t_1)^{r-1} \left[\frac{\partial^r \Phi(t_1, t_2)}{\partial t_2^r} \right]_{t_2=0} dt_1. \tag{14}$$

Since $\tilde{m}$ is bounded, its rth moment exists, For $-\infty < t_1 \leq \epsilon > 0$ and $|t_2| \leq \epsilon > 0$, it can be shown that the moment generating function $\Phi(t_1, t_2)$ exists and is continuous and equal to

$$\Phi(t_1, t_2) = |B|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2\sigma^2} \beta' X' (I - B^{-1}) X \beta \right\}, \tag{15}$$

where $B = (I - 2\sigma^2 t_1) I - 2\sigma^2 t_2 Q$. Applying the formula (14) to the function (15) gives

$$\mathcal{E}(\tilde{m}) = q_1 J\left(1, \frac{T}{2} + 2\right) + \sigma^2 g_1 J\left(1, \frac{T}{2} + 1\right), \tag{16}$$

where $q_1 = \beta' X' Q X \beta$, $g_1 = trQ$ and

$$\begin{aligned} J(r, n) &= \frac{1}{\Gamma(r)} \int_{-\infty}^{0} (-t_1)^{r-1} (1 - 2\sigma^2 t_1)^{-n} \\ &\quad \exp\left\{ -\frac{\beta' X' X \beta}{2\sigma^2} + \frac{\beta' X' X \beta}{2\sigma^2 (1 - 2\sigma^2 t_1)} \right\} dt_1. \end{aligned}$$

It may be noted that the second term on the right-hand side of equation (16) is zero because $trQ = 0$. The integral $J(r, n)$ can be easily evaluated by expanding the integrand into an infinite series (see Sawa 1972, pp. 677-678).

Similarly, we can use the formula (14) to find the second, third, and fourth moments of $\tilde{m}$. In preparation for stating these moments compactly, we need to introduce the following notation : For $i = 1, 2, 3, 4$,

$$q_i = \beta' X' Q^i X \beta;\ g_i = trQ^i;$$
$$f_1 = 12 q_1 q_2;\ f_2 = 6 q_1 g_2 + 24 q_3;\ f_3 = 24 q_1^2 q_2;$$
$$f_4 = 96 q_1 q_3 + 48 q_2^2 + 12 q_1^2 g_2;\ f_5 = 32 q_1 g_3 + 48 q_2 g_2 + 192 q_4;$$

and

$$f_6 = 12 g_2^2 + 48 g_4\,. \tag{17}$$

In terms of this notation

$$\mathcal{E}(\tilde{m}^2) = q_1^2 J\left(2, \frac{T}{2} + 4\right) + 4\sigma^2 q_2 J\left(2, \frac{T}{2} + 3\right)$$

$$+ 2\sigma^4 g_2 J\left(2, \frac{T}{2}+2\right), \tag{18}$$

$$\mathcal{E}(\tilde{m}^3) = \frac{1}{2}\left[q_1^3 J\left(3, \frac{T}{2}+6\right) + \sigma^2 f_1 J\left(3, \frac{T}{2}+5\right) + \sigma^4 f_2 J\left(3, \frac{T}{2}+4\right)\right.$$

$$\left. + 8\sigma^6 g_3 J\left(3, \frac{T}{2}+3\right)\right], \tag{19}$$

and

$$\mathcal{E}(\tilde{m}^4) = \frac{1}{6}\left[q_1^4 J\left(4, \frac{T}{2}+8\right) + \sigma^2 f_3 J\left(4, \frac{T}{2}+7\right) + \sigma^4 f_4 J\left(4, \frac{T}{2}+6\right)\right.$$

$$\left. + \sigma^6 f_5 J\left(4, \frac{T}{2}+5\right) + \sigma^8 f_6 J\left(4, \frac{T}{2}+4\right)\right]. \tag{20}$$

If $\beta = 0$ and $X'X$ is nondiagonal, then real multicollinearity is absent. In this case, it follows equation (16) that $\mathcal{E}(\tilde{m}) = 0$ and from equations (18) - (20) that $\mathcal{E}(\tilde{m}^r) \neq 0$ for $r = 2, 3$, and 4. If, in addition, $\tilde{m}$ is a nonnegative or nonpositive random variable, then the distribution of $\tilde{m}$ is degenerate at 0. Thus, in some cases where real multicollinearity is absent, $\tilde{m}$ takes the correct value 0 with probability 1 but the VIF's give misleading results by taking values greater than 1.

2.7 *Analysis of a general case.* The exact cumulative distribution function (cdf) of $\tilde{m}$ has been numerically evaluated by use of (13) for suitably chosen cases. We used Koerts and Abrahamse's (1969) subroutine called FQUAD to evaluate (13). However, it takes prohibitively long computation time to evaluate the exact cdf to sufficient accuracy when the noncentrality parameters, μ_i's, are large. In these cases, it is difficult to control the accuracy since it is difficult to decide when to truncate the range of the integral in (13) so that the truncation error is negligible. Therefore, this computation was restricted to the following centred model:

$$y = \iota[\beta_0 + (1/T)\iota' Z\delta] + AZD^{*-\frac{1}{2}}\bar{\delta} + u, \tag{21}$$

where the symbols ι, A, Z, and D^* are as explained in equations (4) and (5), $\beta = (\beta_0, \delta')'$ and $\bar{\delta} = D^{*\frac{1}{2}}\delta$. The $\bar{\delta}_j$ denotes the jth element of $\bar{\delta}$. If we treat $\bar{\delta}$ as the vector of unknown coefficients and set $\beta_0 = -(1/T)\iota' Z\delta$, then K and X in Q get replaced by $K-1$ and $AZD^{*-\frac{1}{2}}$, respectively, and the noncentrality parameter of y_i^{*2} in (10) will be equal to the square of the ith element of

$(1/\sigma)P'AZD^{*-\frac{1}{2}}\bar{\delta}$. Since $D^{*-\frac{1}{2}}Z'AZD^{*-\frac{1}{2}}$ is a correlation matrix, the values of noncentrality parameters will not be unmanageably large.

Let us return to Draper and Smith's (1966, p. 366) data used in Section 2.4. From these data the following two correlation matrices for four nonconstant regressors were computed:

$$\begin{pmatrix} 1.000 & .229 & -.824 & -.245 \\ & 1.000 & -.139 & -.973 \\ & & 1.000 & .029 \\ & & & 1.000 \end{pmatrix} \tag{22}$$

and

$$\begin{pmatrix} 1.000 & 0 & -.382 & -.031 \\ & 1.000 & -.478 & -.323 \\ & & 1.000 & .294 \\ & & & 1.000 \end{pmatrix} \tag{23}$$

It can be seen that there are no orthogonal pairs of variables among the regressors whose correlation matrix is (22). Of the four regressors having the matrix (23) as their correlation matrix, the first two are orthogonal to each other. Table 2 and three other tables which are not presented here are based on the correlation matrix (22) and Tables 3 and 4 and two other tables which are not presented here are based on the correlation matrix (23). (The tables which are not presented here are available from the authors on request.) The symbol r_{ij} appearing in these tables refers to the correlation coefficient between the regressors whose coefficients are nonzero. The values of m_L and m_U implied by the matrices (22) and (23) are also presented in these tables.

For each correlation matrix, four pairs of coefficients were taken to have the zero values. For each of these pairs, one of the remaining two coefficients was set equal to 1 and the other was allowed to take 6 different values between -5 and 5. Throughout σ^2 was set equal to 1. The absolute value of the error that is made in truncating the range $(0,\infty)$ of the integral in equation (13) was kept below 0.0000001. Tables 2–4 and the excluded tables contain the values of the mean, mode, median, the variance, the skewness, and the kurtosis of the distribution of $\tilde{m}$ for each case we considered. We employed Simpson's composite rule to numerically evaluate the integral $J(r,n)$ involved in the moments formulas (16) and (18)–(20). The values of d defined in (12) are also presented in these tables.

The properties of $\tilde{m}$ can be inferred from the tables. The first important feature of these tables is that the probability that $m_L \le \tilde{m} \le -1$ is either zero or negligible even when m_L is much smaller than -1. Also, the probability that $\tilde{m}$ is positive is greater than $\frac{1}{2}$ and in many cases markedly so. This result can be easily seen from the tabulated median which is positive in all cases we considered. In fact, each distribution is shifted towards positive values. this asymmetry about 0 generally increases with the magnitude of one of the two

Table 2

Exact CDF of Theil's measure of multicollinearity
$m_L = -2.910,\ m_U = .998,\ r_{24} = -.973, \bar{\delta}_1 = \bar{\delta}_3 = 0,\ \bar{\delta}_2 = 1$

c	$\bar{\delta}_4 = 5$	$\bar{\delta}_4 = 1$	$\bar{\delta}_4 = .1$	$\bar{\delta}_4 = -.1$	$\bar{\delta}_4 = -1$	$\bar{\delta}_4 = -5$	d
1	1.000	1.000	1.000	1.000	1.000	1.000	1
.95	1.000	1.000	1.000	1.000	1.000	.997	.99
.85	.975	1.000	1.000	1.000	.999	.838	.96
.75	.855	.999	.998	.997	.988	.496	.94
.65	.654	.995	.990	.986	.952	.238	.91
.55	.446	.982	.967	.958	.880	.107	.89
.45	.282	.952	.921	.904	.772	.049	.86
.35	.171	.894	.843	.816	.636	.023	.83
.25	.103	.798	.728	.694	.490	.011	.81
.15	.062	.656	.580	.545	.354	.005	.78
.05	.038	.476	.415	.387	.242	.003	.76
−.05	.024	.324	.283	.264	.165	.001	.73
−.15	.015	.236	.205	.191	.117	.001	.71
−.25	.009	.175	.151	.140	.083		.68
−.35	.006	.130	.111	.103	.060		.66
−.45	.004	.097	.082	.076	.043		.63
−.55	.002	.072	.061	.055	.030		.60
−.65	.001	.053	.044	.040	.021		.58
−.75	.001	.039	.032	.029	.015		.55
−.85		.028	.023	.021	.010		.53
−.95		.020	.016	.015	.007		.50
$E(\tilde{m})$	.5350	.0031	.0516	.0747	.2122	.7230	
Median	.557	.063	.101	.122	.257	.751	
Mode	.643	.060	.091	.113	.299	.797	
$var(\tilde{m})$	.0525	.1159	.1163	.1159	.1060	.0207	
γ_3	−1.432	−1.335	−1.258	−1.235	−1.205	−1.556	
γ_4	6.707	6.024	5.718	5.798	5.718	7.292	

Table 3

Exact CDF of Theil's measure of multicollinearity
$m_L = -1.004,\ m_U = .616,\ r_{23} = -.478, \bar{\delta}_1 = \bar{\delta}_4 = 0,\ \bar{\delta}_2 = 1$

c	$\bar{\delta}_4 = 5$	$\bar{\delta}_4 = 1$	$\bar{\delta}_4 = .1$	$\bar{\delta}_4 = -.1$	$\bar{\delta}_4 = -1$	$\bar{\delta}_4 = -5$	d
.65	1.000	1.000	1.000	1.000	1.000	1.000	1.02
.55	1.000	1.000	1.000	1.000	1.000	.996	.96
.45	.986	1.000	1.000	.999	.996	.821	.90
.35	.878	.998	.994	.992	.964	.447	.84
.25	.655	.983	.965	.953	.854	.184	.77
.15	.414	.904	.860	.833	.641	.065	.71
.05	.228	.660	.595	.559	.355	.021	.65
−.05	.111	.302	.236	.216	.127	.006	.59
−.15	.048	.148	.104	.093	.050	.002	.53
−.25	.018	.070	.045	.039	.019		.47
−.35	.006	.030	.018	.015	.007		.40
−.45	.002	.011	.006	.005	.002		.34
−.55		.004	.002	.001	.001		.28
−.65		.001					.22
$E(\tilde{m})$	.1655	−.0106	.0182	.0299	.0987	.3458	
Median	.188	.011	.027	.035	.099	.365	
Mode	.238	.008	.008	.008	.061	.389	
$var(\tilde{m})$	.0283	.0210	.0191	.0193	.0213	.0136	
γ_3	−.6887	−.9394	−.7160	−.6227	−.4199	−.9922	
γ_4	3.445	4.751	4.864	4.726	3.973	4.449	

Table 4

Exact CDF of Theil's measure of multicollinearity
$m_L = -1.004,\ m_U = .616,\ r_{12} = 0, \bar{\delta}_3 = \bar{\delta}_4 = 0,\ \bar{\delta}_1 = 1$

c	$\bar{\delta}_2 = 5$	$\bar{\delta}_2 = 1$	$\bar{\delta}_2 = .1$	$\bar{\delta}_2 = -.1$	$\bar{\delta}_2 = -1$	$\bar{\delta}_2 = -5$	d
.65	1.000	1.000	1.000	1.000	1.000	1.000	1.02
.55	.999	1.000	1.000	1.000	1.000	1.000	.96
.45	.946	.999	1.000	1.000	1.000	.985	.90
.35	.733	.987	.997	.998	.997	.811	.84
.25	.459	.933	.976	.980	.973	.464	.77
.15	.244	.793	.885	.892	.852	.200	.71
.05	.114	.530	.620	.623	.539	.074	.65
−.05	.047	.232	.237	.232	.188	.024	.59
−.15	.017	.109	.101	.097	.074	.007	.53
−.25	.005	.049	.042	.040	.028	.002	.47
−.35	.001	.020	.016	.015	.010		.40
−.45		.007	.006	.005	.003		.34
−.55		.002	.002	.001	.001		.28
−.65							.22
$E(\tilde{m})$	.2433	.0341	.0124	.0124	.0341	.2434	
Median	.266	.041	.022	.023	.041	.261	
Mode	.316	.008	.008	.008	.008	.291	
$var(\tilde{m})$	.0236	.0238	.0172	.0162	.0153	.0153	
γ_3	−.7947	−.6057	−.8033	−.8457	−.7893	−.8628	
γ_4	3.695	4.286	5.134	5.256	5.133	4.176	

nonzero coefficients. For some combinations of parameter values, such as ($m_U =$ .998, $\bar{\delta}_1 = 0, \bar{\delta}_2 = 1, \bar{\delta}_3 = 0, \bar{\delta}_4 = -5$, $r_{24} = -.973$) or ($m_U = .998, \bar{\delta}_1 = 1, \bar{\delta}_2 = 0, \bar{\delta}_3 = -5, \bar{\delta}_4 = 0$, $r_{13} = -.824$), the probability that $\tilde{m}$ is positive is greater than .99. It is also true that the probability that $\tilde{m}$ takes values in the interval $(-.05, .05)$ increases as at least one of the two nonzero coefficients moves closer to zero. Since real multicollinearity arises when at least two coefficients corresponding to the nonorthogonal columns of Z are nonzero, these results show that the probability that $\tilde{m}$ takes values outside the interval $(-.05, .05)$ increases as the degree of real multicollinearity increases. This is a desirable property of $\tilde{m}$.

The second important feature is that with three exceptions the spread of the distribution of $\tilde{m}$ decreases as one of the two nonzero coefficients moves in either direction away from zero. This property is demonstrated by the coefficient of variation which is not tabulated but can be calculated by dividing the square root of each tabulated variance by the corresponding tabulated mean. It will be noted that this relationship is exactly opposite to that of the preceding paragraph. That is, asymmetry towards positive values is associated with a small spread; the smaller spread means the distribution is tighter and centred at positive values when real multicollinearity is present. The tables also show that the speed with which the tightness of the distribution of $\tilde{m}$ increases as the magnitude of a coefficient increases depends on the signs of r_{ij} and the corresponding coefficients. If we compare $\tilde{m}$ with the square roots of the VIF's, the former would be certainly preferred in terms of the ability to detect real multicollinearity, particularly when the absolute values of m_U and coefficients are large.

It is possible that when m_U is much smaller than 1, m_L is not much smaller than -1. The above properties of the distribution of $\tilde{m}$ are more clearly noticeable when m_U is very close to 1 than when it is not close to 1. For example, when $m_U = .616$, the variance of $\tilde{m}$ is not very different for different values of coefficients and r_{ij} we considered. Also, in these cases, the asymmetry of the distribution of $\tilde{m}$ towards positive values is not very pronounced. When there are more than two nonconstant regressors in the model, the distribution of $\tilde{m}$ does not seem to be as sensitive to changes in the value of r_{ij} as it is to changes in the values of m_U and coefficients. These results cast doubt on the ability of the regressors' correlations to correctly diagnose real multicollinearity.

Table 4 covers the cases with $r_{12} = 0$. In these cases, there is no real multicollinearity because all the coefficients corresponding to the nonorthogonal columns of Z are zero. The nondegenerate distribution of $\tilde{m}$ tabulated in Table 4 for these case shows that when there is no real multicollinearity, $\tilde{m}$ can take values outside the interval $(-.05, .05)$ with positive probability, though this probability decreases with the magnitude of at least one of the two nonzero coefficients. This shows that $\tilde{m}$ does not always give an accurate estimate of the degree of real multicollinearity actually present in model (1).

Since we have already discussed the spread of the distribution of $\tilde{m}$, let us now consider its moments other than its variance. The mean of $\tilde{m}$ is positive in all but two cases we considered. It is generally higher for the coefficients of bigger magnitude. It cannot adequately represent the degree of real multicollinearity actually present in model (1), except when $\beta = 0$ for the following reasons: (i) The distribution of $\tilde{m}$ is asymmetric and (ii) for the parameter combination, $(m_U = .998, \bar{\delta}_1 = 0, \bar{\delta}_2 = 1, \bar{\delta}_3 = 0, \bar{\delta}_4 = 1, r_{24} = -.973)$, which represents a high degree of real multicollinearity, $\mathcal{E}(\tilde{m})$ is as small as .0031 (Table 2) and for the parameter combination, $(m_U = .616, \bar{\delta}_1 = 1, \bar{\delta}_2 = -5, \bar{\delta}_3 = 0, \bar{\delta}_4 = 0, r_{12} = 0)$, which represents the zero degree of real multicollinearity, $\mathcal{E}(\tilde{m})$ is as big as .2434 (Table 4). The median and modal values of $\tilde{m}$ are positive in all cases we considered. With one exception, the mean is smaller than the median. However, the modal value is smaller than the mean in several cases when $m_U = .616$. In some cases with $m_U = .616$ the mode of the distribution of $\tilde{m}$, unlike its median or mean, is relatively insensitive to changes in the value of a coefficient. In several cases we considered, the median or mode increases as a coefficient moves away from zero. These results show that the absolute value of $\tilde{m}$ is a better indicator of the actual degree of real multicollinearity than its mean, mode, or median. In cases, where m_U is sufficiently close to 1, the larger the absolute value of $\tilde{m}$, the greater the degree of real multicollinearity.

The distributions are also skewed; this property can be seen from the tabulated coefficient of skewness (γ_3) which is negative in all cases we considered. This means that the long tails of these distributions are on the negative side. The γ_3 is smaller in magnitude when $m_U = .616$ than when $m_U = .998$. The values of the coefficient of kurtosis (γ_4) presented in the tables indicate that the probability density curves of $\tilde{m}$ are taller and slimmer than a normal curve in the neighbourhood of their modes. This relative peakedness is more pronounced when $m_U = .998$ than when $m_U = .616$.

On the basis of the values of c and d given in the first and the last columns of the tables, respectively, we could make some comparative statements between $\hat{m}$ and $\tilde{m}$ defined in (11). The variable $\hat{m}$ has the disadvantage that it transforms the small absolute values $\tilde{m}$ takes with positive probability in the absence of real multicollinearity into large positive values.

5. Conclusion

We have found that though Theil's (1971) measure of mutlicollinearity has its defects, it is preferable to other measures used in the statistics literature to diagnose multicollinearity. More specifically, Theil's measure, unlike other multicollinearity diagnostics adheres to economists' definition of near multicollinearity and can differentiate between real and apparent multicollinearities in numerous cases. To confirm this claim, we have given numerical tables of the exact cdf of Theil's measure in the normal case. A by-product of the confirmation is the derivation of the first four moments of Theil's measure.

Acknowledgement. Thanks are due to a referee for helpful comments.

REFERENCES

DRAPER, N. R. and SMITH, H. (1966). *Applied Regression Analysis.* Wiley, New York.

IMHOF, P. J. (1961) Computing the distribution of quadratic forms in normal variates, *Biometrika*, **68**, 419–426.

KOERTS, J. and ABRAHAMSE, A. P. J. (1969). *On the theory and Application of the General Linear Model*, University of Rotterdam Press, Rotterdam.

MEHTA, J. S. and SWAMY, P. A. V. B. (1978). The existence of moments of some simple Bayes estimators of coefficients in a simultaneous equation model. *J. Econometrics*, **7**, 1–13.

RAO, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd edition, Wiley, New York.

SAWA, T. (1972). Finite-sample properties of the *k*-class estimators. *Econometrica*, **40**, 653–680.

STEWART, G. W. (1987). Collinearity and least squares regression (with discussion). *Statist. Sci.*, **2** , 68–100.

SWAMY. P. A. V. B., MEHTA, J. S., THURMAN, S.S., and IYENGAR, N. S. (1985). A generalized multicollinearity index for regression analysis. *Sankhyā*, *Ser. B*, **47**, 401–431.

THEIL, H. (1971). *Principles of Econometrics*, Wiley, New York.

DEPT. OF MATHEMATICS
TEMPLE UNIVERSITY
PHILADELPHIA, PA 19122
U. S. A.

MONETARY & FINANCIAL STUDIES SEC.
FEDERAL RESERVE BOARD
20TH AND CONSTITUTION AVE. N.W.
WASHINGTON, DC 20557
U. S. A.

257, I CROSS, II BLOCK
BANASHANKARI III STAGE
BANGALORE 560085
INDIA