Journal of Multivariate Analysis 126 (2014) 137–152

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Coefficient of determination for multiple measurement error models☆

C.-L. Cheng [a], Shalabh [b,*], G. Garg [c]

[a] Institute of Statistical Science, Academia Sinica, Taipei, Taiwan, ROC

[b] Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur - 208 016, India

[c] Decision Sciences Area, Indian Institute of Management Lucknow, Lucknow - 226 013, India

ARTICLE INFO

Article history:
Received 27 December 2012
Available online 5 February 2014

AMS subject classifications:
62J05
62H12

Keywords:
Measurement error
Linear regression
Coefficient of determination (R^2)
Ultrastructural model
Non-normal distribution

ABSTRACT

The coefficient of determination (R^2) is used for judging the goodness of fit in a linear regression model. It is the square of the multiple correlation coefficient between the study and explanatory variables based on the sample values. It gives valid results only when the observations are correctly observed without any measurement error. The conventional R^2 provides invalid results in the presence of measurement errors in the data because the sample R^2 becomes an inconsistent estimator of its population counterpart which is the square of the population multiple correlation coefficient between the study and explanatory variables. The goodness of fit statistics based on the variants of R^2 for multiple measurement error models have been proposed in this paper. These variants are based on the utilization of the two forms of additional information from outside the sample. The two forms are the known covariance matrix of measurement errors associated with the explanatory variables and the known reliability matrix associated with the explanatory variables. The asymptotic properties of the conventional R^2 and the proposed variants of R^2 like goodness of fit statistics have been studied analytically and numerically.

1. Introduction

The linear regression analysis has a prominent role in extracting the statistical information from the data through the determination of relationship between the study and explanatory variables. An adequate linear regression model provides valid statistical inferences on various applications including the forecasting. The success of linear regression analysis lies on the adequacy of the fitted model in explaining the variations in the data set. A popular tool to determine the adequacy of the fitted model is the coefficient of determination and the adjusted version. The coefficient of determination is popularly known as R^2 and its adjusted version is called as adjusted R^2. They are treated as summary measures for the goodness of fit of any linear regression model. The R^2 is based on the proportion of variability of the study variable that can be explained through the knowledge of a given set of explanatory variables. It is the square of the multiple correlation coefficient between the study variable and all the explanatory variables present in the linear regression model. The R^2 and its adjusted version are also used for the model selection. For example, if there are several fitted models available from the same data set, then

☆ The authors are grateful to the referees for their comments to improve the exposition of the paper.

* Corresponding author.
E-mail addresses: clcheng@stat.sinica.edu.tw (C.-L. Cheng), shalab@iitk.ac.in, shalabh1@yahoo.com (Shalabh), ggarg@iiml.ac.in, ggarg31@gmail.com (G. Garg).

http://dx.doi.org/10.1016/j.jmva.2014.01.006
0047-259X/

a model with the least lack of fit is preferred and can be determined based on the values of the coefficient of determination or its adjusted version. Although R^2 and its adjusted versions have certain limitations, see [12], but in spite of them, they remain a popular choice among practitioners.

The research work in obtaining the different suitable forms of the coefficient of determination for various situations has been addressed in the literature by several researchers. Eshima and Tabata [6,7] proposed the coefficient of determination in entropy form for generalized linear models. Renaud and Victoria-Feser [28] presented a robust coefficient of determination in regression. Tjur [42] proposed a coefficient of determination for the logistic regression model, see also [14,18]. Huang and Chen [16] addressed the issue of the coefficient of determination in the local polynomial model. Hössjer [15] discussed the role of the coefficient of determination in the mixed regression model. Linde and Tutz [44] considered the coefficient of determination in the case of association in a regression framework. Srivastava and Shobhit [39] proposed a family of coefficients of determination in the linear regression model. Marchand [21] discussed the point estimation of the coefficient of determination, see also [22]. Lipsitz et al. [19] discussed the partial correlation coefficient and the coefficient of determination for the multivariate normal repeated measures data. Tanaka and Huba [41] presented a general coefficient of determination for the covariance structure models under arbitrary generalized least squares estimation. Nagelkerke [24] presented a generalization of the coefficient of determination. McKean and Sievers [23] obtained a new coefficient of determination for the least absolute deviation analysis. Knight [17] and Hilliard and Lloyd [13] discussed the role of the coefficient of determination in the simultaneous equation models. Ohtani [25] derived the density of R^2 and its adjusted version. He also analyzed their risk performance under an asymmetric loss function in the misspecified linear regression model.

One of the fundamental assumptions in using the coefficient of determination in the linear regression analysis is that all the observations on the study and explanatory variables are correctly observed. Many times in practical situations, the variables are not correctly observable and the measurement errors creep into the data. If the magnitude of measurement errors is negligible, then it may not pose any big challenge to the derived statistical inferences. On the other hand, when the magnitude of measurement errors is large, then it disturbs the optimal properties of the estimators. A serious consequence of measurement errors in linear regression analysis is that the ordinary least squares estimator (OLSE) of the regression coefficients becomes biased and inconsistent. Note that the same OLSE is the best linear unbiased estimator of the regression coefficients in the absence of measurement errors in the data. The coefficient of determination (R^2) is a function of OLSE. So consequently, the presence of the measurement error disturbs the properties of R^2. The value of R^2 obtained by ignoring the measurement errors becomes misleading and may provide incorrect statistical inferences. So we are faced with the question of how to judge the goodness of fit in the linear regression model when the observations are contaminated with measurement errors. Such an issue has never been addressed in the literature, to the best of our knowledge.

It may also be noted that the expression of conventional R^2 in the multiple linear regression model is based on the analysis of variance. It is defined as the ratio of the sum of squares due to regression and the total sum of squares. Unfortunately, such analysis of variance in the setup of measurement error models is not possible. This is due to the nonexistence of the moments of the estimators and the complicated structure of moments, if they exist in some cases, see [2,1]. So the only option left is possibly to look at the structure of R^2 and adjust it in the framework of the measurement error model so as to reflect the goodness of fit. We have attempted in this direction.

In order to obtain the consistent estimators of regression coefficients in the presence of measurement errors in the data, the OLSE is adjusted for its inconsistency. Such an adjustment is done by using the additional information from outside the sample. Various forms of additional information can be used to obtain the consistent estimators, see [3,8] etc. for more details. In the context of the multiple measurement error model, there are two possible forms of additional information which can be used to obtain consistent estimators of the regression coefficient vector. These two forms are based on the knowledge of the covariance matrix of measurement errors associated with explanatory variables and the knowledge of the reliability matrix of explanatory variables, see, e.g. [3,29,9,10,31–33] etc. Since the form of the conventional R^2 is directly related to OLSE of the regression coefficient in the no-measurement error linear regression model, an idea to obtain statistics for judging the goodness of fit in the measurement error model can be based on the form of conventional R^2. Our objective in this paper is to use both types of available information and obtain an appropriate form of the coefficient of determination which can be used to judge the goodness of a fit in the measurement error models.

The plan of the paper is as follows. The multivariate ultrastructural model and the various statistical assumptions are described in Section 2. In Section 3, we demonstrate the inconsistency of the coefficient of determination under the ultrastructural form of the measurement error model. We propose two goodness of fit statistics based on R^2-like expressions. These statistics are consistent for the population counterpart of R^2 which is the square of the population multiple correlation coefficient. The asymptotic distributions of the proposed R^2 like goodness of fit statistics are derived under the specification of the ultrastructural measurement error model in Section 5. In order to study the small sample properties of the proposed goodness of fit statistics, Monte Carlo simulation experiments are conducted. The findings of the simulation study are presented in Section 6 followed by some concluding remarks in Section 7. Lastly, the proof of the results is given in the Appendix.

2. The model

We consider the following exact relationship between the $(n \times 1)$ vector of values of study variable η and the $(n \times p)$ matrix Ξ of n values on each of the p explanatory variables:

$$\eta = \alpha \mathbf{1}_n + \Xi \beta, \tag{2.1}$$

where α is the intercept term, $\mathbf{1}_n$ is the $(n \times 1)$ vector of elements unity (1's), and β is the $(p \times 1)$ vector of regression coefficients.

When the observations on the study and explanatory variables are contaminated with measurement errors, then η and Ξ cannot be accurately observed. We assume that they are observed with additive measurement errors as

$$y = \eta + \epsilon \tag{2.2}$$

and

$$X = \Xi + \Delta, \tag{2.3}$$

respectively. Here, $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)'$ is the $(n \times 1)$ vector of measurement errors associated with η and $\Delta = (\delta_1, \delta_2, \ldots, \delta_n)'$ is the $(n \times p)$ matrix of measurement errors associated with the explanatory variables in Ξ, respectively. Further, we assume that the true values of explanatory variables are expressible as

$$\Xi = M + \Phi, \tag{2.4}$$

where $M = E(\Xi) = (\mu_1, \mu_2, \ldots, \mu_n)'$ is the $(n \times p)$ matrix of unknown means (constants) of true explanatory variables in Ξ and $\Phi = (\phi_1, \phi_2, \ldots, \phi_n)'$ is the random matrix associated with Ξ. The $(p \times 1)$ random vectors $\phi_1, \phi_2, \ldots, \phi_n$ of matrix Φ are assumed to be independently and identically distributed with mean vector 0, covariance matrix Σ_ϕ and finite fourth moment. We also assume that the $(p \times 1)$ random vectors $\delta_1, \delta_2, \ldots, \delta_n$ of matrix Δ are assumed to be independently and identically distributed with mean vector 0, covariance matrix Σ_δ and finite moments up to order four. Further, the random variables $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ are assumed to be independently and identically distributed with mean zero and variance σ_ϵ^2 and finite fourth moment.

Eqs. (2.1)–(2.4) describe the setup of an ultrastructural model (see [5]) which is the synthesis of the structural form and the functional form of the measurement error model, see [11,3,34–37]. The two forms of the measurement error models, viz., the structural and the functional as well as the classical regression model with no measurement errors, can be obtained as particular cases. When all the row vectors of M are identical, implying that the rows of X are identically and independently distributed, then we get the specification of a structural model. When Σ_ϕ is equal to a null matrix, implying that Φ is identically equal to the null matrix and consequently the matrix X is fixed but is measured with error, then we obtain the specification of a functional model. When both the matrices Σ_ϕ and Σ_δ are equal to the null matrix, implying that both Φ and Δ are identically equal to the null matrix, then the X matrix becomes fixed and is measured without any measurement error. In such a situation, we get the specification of the classical regression model. Thus the ultrastructural model facilitates the study of the functional model, the structural model as well as the classical regression model under the same framework.

We further assume that $\lim_{n\to\infty} n^{-1}M'PM =: \Sigma_\mu$ which is a symmetric and positive definite matrix where $P = I_n - n^{-1}\mathbf{1}_n\mathbf{1}_n'$. This assumption is needed for the validity of asymptotic results and avoids the possibility of any trend in the data, see [30].

3. Coefficient of determination in the classical regression model

First we state the definition of convergence in probability and notations used for better understanding.

Definition. Let $\{Z_n\}$ be a sequence of random variables defined on some probability space. Then $\{Z_n\}$ is said to converge to the random variable Z if for every $\vartheta > 0$,

$$P\{|Z_n - Z| > \vartheta\} \to 0 \quad \text{as } n \to \infty.$$

We say that $\{Z_n\}$ converges to Z on probability as n goes to infinity and is denoted as $\operatorname{plim}_{n\to\infty} Z_n = Z$. The notation plim denotes the "probability in limit". The probability in limit also indicates the consistency property of an estimator. If $\hat{\varpi}$ is a consistent estimator of ϖ, then $\operatorname{plim}_{n\to\infty} \hat{\varpi} = \varpi$.

Next we discuss the performance of R^2 in the classical regression model under the usual assumptions. Let us consider the following classical multiple linear regression model where explanatory variables are non-stochastic and measurement errors are absent:

$$y^* = \alpha\mathbf{1} + X^*\beta + u,$$

where y^* is the $(n \times 1)$ vector of values on the study variable, X^* is the $(n \times p)$ matrix of n values on each of the p non-stochastic explanatory variables, α is the intercept term, β is the $(p \times 1)$ vector of regression slopes, and u is the $(n \times 1)$ vector of disturbances. The notation of $*$ denotes that the concerned variables are obtained without any measurement errors. Under this classical multiple linear regression model, the coefficient of determination is defined as

$$\begin{aligned} R^{*2} &= \frac{b^{*\prime}X^{*\prime}PX^*b^*}{y^{*\prime}Py^*} \\ &= \frac{y^{*\prime}PX^*(X^{*\prime}PX^*)^{-1}X^{*\prime}Py^*}{y^{*\prime}Py^*}, \end{aligned} \tag{3.1}$$

where $b^* = (X^{*\prime}PX^*)^{-1}X^{*\prime}Py^*$ is the ordinary least squares estimator β. Assuming $E(u) = 0$ and covariance matrix $V(u) = \sigma^2 I$ with other usual assumptions of the classical linear regression analysis, it can be easily shown that

$$\operatorname*{plim}_{n\to\infty}(R^{*2} - \theta^*) = 0,$$

where

$$\theta^* = \frac{\beta'(n^{-1}X^{*\prime}PX^*)\beta}{\beta'(n^{-1}X^{*\prime}PX^*)\beta + \sigma^2}$$

denotes the square of the population multiple correlation coefficient between the study and explanatory variables and is the population counterpart of R^2. Thus it is established that R^{*2} is a consistent estimator of θ^*. This conclusion remains true only in the absence of measurement errors in the data. It may be noted that R^{*2} is a biased estimator of θ^*.

A systematic study of the properties of R^2 and its adjusted version under the normality of disturbances was conducted by Cramer [4]. Ohtani and Giles [26] studied the relative performance of R^2 and its adjusted version with respect to the criterion of risk under an absolute error loss function. Relaxing the assumption of normal distribution and considering the multivariate t-distribution of errors, Ohtani and Hasegawa [27] analyzed the properties of R^2 and its adjusted version in the presence of the specification error. The specification error relates to the mis-specification of the multiple linear regression model with respect to the explanatory variables in the sense that either some important explanatory variables are not included or some unimportant explanatory variables are included in the model. In all these studies, the exact expressions of the properties of R^2 and/or its adjusted version turn out to be quite intricate. Consequently, no clear and useful conclusions can be drawn. These expressions are therefore evaluated numerically for some selected values of the parameters and the sample sizes. Obviously, the observations emerging from such numerical exercises are limited in their scope and value.

An alternative route is to follow some asymptotic theory and use it to obtain the suitable approximations for the exact expressions. Generally, the approximations obtained in this way are much simpler in comparison to their exact counterparts and it is not difficult to draw meaningful inferences from them. In fact, the application of asymptotic theory in many situations does not require the assumption of any specific distribution of errors like normal. It may then be sufficient to assume the finiteness of first few moments of the distribution and thus it is possible to draw fairly general conclusions.

Ullah and Srivastava [43] obtained the approximations for the exact moments of R^2 by employing the small disturbance asymptotic theory. The conclusions drawn from such approximations have their validity only when θ^* is near 1, i.e., the case of perfect fit. Thus the use of small disturbance asymptotic theory in analyzing the performance of R^2 does not seem to be appropriate and worthwhile; see also [38] who examined the closeness of small disturbance and other asymptotic approximations to the exact results. Srivastava et al. [40] utilized the large sample asymptotic theory and derived the approximations for the biases and mean squared errors of R^2 and its adjusted version.

Now we illustrate the performance of the coefficient of determination in the measurement error model. Note that Eqs. (2.1)–(2.2) can jointly be written as

$$y = \alpha\mathbf{1}_n + \Xi\beta + \epsilon. \tag{3.2}$$

Under the model (3.2), the coefficient of determination can be defined along similar lines as in the case of the classical regression model without measurement errors as

$$R^{\dagger 2} = \frac{y'P\Xi(\Xi'P\Xi)^{-1}\Xi'Py}{y'Py}. \tag{3.3}$$

Let $\Sigma^\dagger := \operatorname{plim}_{n\to\infty} n^{-1}\Xi'P\Xi$, then

$$\operatorname*{plim}_{n\to\infty}(R^{\dagger 2} - \theta^\dagger) = 0,$$

where

$$\theta^\dagger = \frac{\beta'\Sigma^\dagger\beta}{\beta'\Sigma^\dagger\beta + \sigma_\epsilon^2}$$

is the square of the population multiple correlation coefficient. So this establishes a sort of similarity between the coefficients of determination in the linear regression models under two cases, viz., with and without measurement errors.

4. Goodness of fit statistics in the measurement error model

The matrix Ξ is not observable in the measurement error model. The values of explanatory variables can only be observed as X with measurement errors given by Δ. So we replace the unobservable Ξ by observable X in (3.3) and attempt to obtain the expression for the goodness of fit statistics based on the form of the coefficient of determination as follows:

$$R^2 = \frac{y'PX(X'PX)^{-1}X'Py}{y'Py}. \tag{4.1}$$

In case, X has no measurement errors, then the coefficient of determination defined in (4.1) is consistent for estimating the parameter

$$\theta = \frac{\beta' \Sigma \beta}{\beta' \Sigma \beta + \sigma_\epsilon^2}, \tag{4.2}$$

where $\Sigma := \text{plim}_{n\to\infty} n^{-1} X'PX$. The ordinary least squares estimate of the regression coefficient becomes inconsistent in the presence of measurement errors in the data. Several approaches are available to find the consistent estimators of regression coefficients in the multiple measurement error models, e.g. use of additional information from outside the sample, instrumental variable method etc. We use the approach based on the use of additional information from outside the sample. There are two popular forms of information which can be used. These forms are the known covariance matrix of measurement errors associated with the explanatory variables and the known reliability matrix associated with the explanatory variables.

First we investigate the consistency of conventional R^2 for the parameter θ in the presence of measurement errors in the model. We present some results in Lemma 1 which are useful in establishing the inconsistency of conventional R^2 defined in (4.1).

Lemma 1. *Under the model (2.1)–(2.4) and the assumptions made in Section 2, we have the following results.*

(i) $\text{plim}_{n\to\infty} n^{-1} X'PX = \Sigma_\mu + \Sigma_\phi + \Sigma_\delta = \Sigma$,
(ii) $\text{plim}_{n\to\infty} n^{-1} X'Py = (\Sigma - \Sigma_\delta)\beta$,
(iii) $\text{plim}_{n\to\infty} n^{-1} y'Py = \beta'(\Sigma - \Sigma_\delta)\beta + \sigma_\epsilon^2$,
(iv) $\lim_{n\to\infty} \Sigma_x = \Sigma$, *where* $\Sigma_x = n^{-1} M'PM + \Sigma_\phi + \Sigma_\delta$.

Proof of the lemma is omitted.

Theorem 1. *Under the results of Lemma 1, the probability in limit of conventional R^2 under the ultrastructural model (2.1)–(2.4) is*

$$\text{plim}_{n\to\infty} R^2 = \frac{\beta'(\Sigma - \Sigma_\delta)\Sigma^{-1}(\Sigma - \Sigma_\delta)\beta}{\beta'(\Sigma - \Sigma_\delta)\beta + \sigma_\epsilon^2}.$$

Clearly, under the measurement error models, $\text{plim}_{n\to\infty} R^2 \neq \theta$, in general. This proves that the R^2 is an inconsistent estimator of θ under this setup.

Consequently, the conclusions obtained by using the conventional R^2 to judge the goodness of the fitted linear regression model will be misleading if the data has measurement errors. Therefore it is not advisable to use the conventional coefficient of determination (R^2) as a tool to decide about the goodness of fit in the measurement error models.

In order to provide a goodness of fit statistic based on the structure of the coefficient of determination, we would like to develop the statistic which is at least a consistent estimator of θ in the presence of measurement errors in the model. The property of unbiasedness is difficult to meet. To fulfill such a requirement, we need some additional information from outside the sample. Such additional information is the same as that is required to estimate β consistently. We consider here two cases. In the first case, we assume that the common covariance matrix of measurement error vectors δ_i, $(i = 1, 2, \ldots, n)$ is known. In the second case, we assume that the reliability matrix of the explanatory variables is known. Such additional information can be available from various resources like the past experience of the researcher, from some similar kind of studies done in the past, some pilot survey etc.

In case such additional information is not available but if the repeated data are available, then such information can be estimated from the sample itself. The forms of the proposed goodness of fit statistics remain the same under the repeated multiple measurement error model and they can be obtained just by replacing Σ_δ or K_x by their respective estimated values. The asymptotic distributions of the statistics thus obtained will be different than those reported in this paper. Deriving these expressions is out of the purview of this paper.

4.1. Σ_δ is known

We attempt to obtain a goodness of fit statistic based on the form of the coefficient of determination which should be at least consistent for estimating the parametric function θ as in (4.2). To do so, we look for the consistent estimators of $\beta' \Sigma \beta$ and σ_ϵ^2. Then replace $\beta' \Sigma \beta$ and σ_ϵ^2 by their respective consistent estimators in θ. Such derived statistic will be a R^2-like statistic for estimating θ. This can be used to check the goodness of fit and considered like the sample R^2 for the measurement error model.

When the covariance matrix Σ_δ is known, the consistent estimator of β in the ultrastructural model (2.1)-(2.4) is given by

$$b_\delta = (S - \Sigma_\delta)^{-1} Sb, \tag{4.3}$$

where $S = n^{-1}X'PX$ and

$$b = (X'PX)^{-1}X'Py \tag{4.4}$$

is the OLSE of β, see [2,8]. We now present Lemma 2 which is a direct consequence of the fact that $\operatorname{plim}_{n\to\infty} b_\delta = \beta$ and Lemma 1:

Lemma 2.

$$\operatorname*{plim}_{n\to\infty}\{n^{-1}y'Py - b_\delta'(S - \Sigma_\delta)b_\delta\} = \sigma_\epsilon^2.$$

Using Lemmas 1(i), 2 and b_δ in (4.3), we propose a new goodness of fit statistic based on the form of the coefficient of determination under the knowledge of Σ_δ as

$$\begin{aligned} R_\delta^2 &= \frac{b_\delta' S b_\delta}{b_\delta' S b_\delta + \{n^{-1}y'Py - b_\delta'(S - \Sigma_\delta)b_\delta\}} \\ &= \frac{b_\delta' S b_\delta}{n^{-1}y'Py + b_\delta'\Sigma_\delta b_\delta}, \quad 0 \le R_\delta^2 \le 1, \end{aligned} \tag{4.5}$$

provided $b_\delta' S b_\delta \ge n^{-1}y'Py + b_\delta'\Sigma_\delta b_\delta$. In case $b_\delta' S b_\delta < n^{-1}y'Py + b_\delta'\Sigma_\delta b_\delta$, we take the value of R_δ^2 as 1.

So the modified coefficient of determination like statistic under the assumption of known Σ_δ can possibly be defined as

$$R_\delta^2 = \min\left(\frac{b_\delta' S b_\delta}{n^{-1}y'Py + b_\delta'\Sigma_\delta b_\delta}, 1\right). \tag{4.6}$$

It can be seen that

$$\operatorname*{plim}_{n\to\infty} R_\delta^2 = \theta.$$

Thus the proposed goodness of fit statistic R_δ^2 can be used to judge the goodness of fit in the linear measurement error model in place of the traditional R^2.

4.2. Reliability matrix is known

The reliability ratio is defined as the ratio of the variances of true and observed values of the explanatory variable. The reliability matrix is the multivariate generalization of reliability ratios. The reliability matrix is defined as

$$K_x = \Sigma_x^{-1}(\Sigma_x - \Sigma_\delta).$$

When K_x is known, then the consistent estimator of β in ultrastructural model (2.1)–(2.4) is given by

$$b_k = K_x^{-1}b, \tag{4.7}$$

see [2,8]. This estimator has its own advantages, see [9,10] for more details. For example, this estimator can be obtained by obtaining the OLSE. We now present a lemma which is a direct consequence of the result $\operatorname{plim}_{n\to\infty} b_k = \beta$ and Lemma 1:

Lemma 3.

$$\operatorname*{plim}_{n\to\infty}\{n^{-1}y'Py - b_k' S K_x b_k\} = \sigma_\epsilon^2.$$

Using Lemmas 1(i), 3 and b_k in (4.7), we propose a new goodness of fit statistic based on the form of the coefficient of determination under the knowledge of K_x, as

$$\begin{aligned} R_k^2 &= \frac{b_k' S b_k}{b_k' S b_k + \{n^{-1}y'Py - b_k' S K_x b_k\}} \\ &= \frac{b_k' S b_k}{n^{-1}y'Py + b_k' S(I_p - K_x)b_k}; \quad 0 \le R_k^2 \le 1, \end{aligned} \tag{4.8}$$

provided $b_k' S b_k \ge n^{-1}y'Py + b_k' S(I_p - K_x)b_k$. In case $b_k' S b_k < n^{-1}y'Py + b_k' S(I_p - K_x)b_k$, we take the value of R_k^2 as 1. So the modified coefficient of determination like statistic under the assumption of the known reliability matrix can be defined as

$$R_k^2 = \min\left(\frac{b_k' S b_k}{n^{-1}y'Py + b_k' S(I_p - K_x)b_k}, 1\right). \tag{4.9}$$

Using the results of Lemma 1 and the consistency of b_k for estimating β, it can be proved that $\operatorname{plim}_{n\to\infty} R_k^2 = \theta$. Thus, it is clear that the proposed coefficient of determination like statistic R_k^2 is a better choice as a measure of goodness of fit in the linear measurement error model in place of the conventional R^2.

4.3. Relation between the estimated reliability matrix and known Σ_δ cases

When K_x is known, then β is consistently estimated by b_k in (4.7). The case of the known reliability matrix in obtaining the consistent estimates of the regression coefficient received more attention after the work of [9]. He suggested that if the reliability matrix is not known from outside the sample and somehow if it is possible to estimate it as $\hat{K}_x$, then the consistent estimator of β is given by

$$\hat{b}_k = \hat{K}_x^{-1} b \tag{4.10}$$

where b in (4.4) is the ordinary least squares estimator of β based on the measurement error ridden observed values of the study and explanatory variables. An interesting observation arises which gives a one-to-one relationship between the two cases of $\hat{K}_x$ and known Σ_δ as follows.

If Σ_δ is known, then

$$b_\delta = (S - \Sigma_\delta)^{-1} S b$$

where Σ_x is estimated by $\hat{\Sigma}_x = \frac{1}{n} X'PX = S$. Then $K_x = \Sigma_x^{-1}(\Sigma_x - \Sigma_\delta)$ can be estimated as

$$\begin{aligned} \hat{K}_x &= S^{-1}(S - \Sigma_\delta) \\ &= I - S^{-1}\Sigma_\delta. \end{aligned} \tag{4.11}$$

In such a case

$$\begin{aligned} b_k &= K_x^{-1} b \\ \Rightarrow \hat{b}_k &= \hat{K}_x^{-1} b \\ &= (S - \Sigma_\delta)^{-1} S b \\ &= b_\delta. \end{aligned} \tag{4.12}$$

If K_x is known

$$\begin{aligned} K_x &= \Sigma_x^{-1}(\Sigma_x - \Sigma_\delta) \\ &= I_p - \Sigma_x^{-1}\Sigma_\delta \\ \Rightarrow \Sigma_x^{-1}\Sigma_\delta &= I_p - K_x \end{aligned}$$

then Σ_δ can be estimated by

$$\begin{aligned} &S^{-1}\hat{\Sigma}_\delta = I_p - K_x \\ &\Rightarrow \hat{\Sigma}_\delta = S - SK_x. \end{aligned}$$

Thus

$$\begin{aligned} \hat{b}_\delta &= (S - \hat{\Sigma}_\delta)^{-1} S b \\ &= (SK_x)^{-1} S b \\ &= K_x^{-1} b \\ &= b_k. \end{aligned} \tag{4.13}$$

The relationships (4.12) and (4.13) indicate that if either Σ_δ is known or K_x is estimated, then both the estimators, viz., b_δ or b_k can be determined from each other.

In our case, if Σ_δ is known or K_x is estimated, then the corresponding coefficient of determination can be determined directly. Moreover, we can also conclude that both R_δ^2 and R_k^2 converge to θ. So the large sample behavior of R_δ^2 and R_k^2 may be similar but their finite sample properties may differ. Moreover, if $\Sigma_\delta = \sigma_\delta^2 I$, then the problem reduces to simply knowing the value of σ_δ^2.

Also, the behavior of values of the coefficient of determination in the two cases, viz., known K_x or estimated K_x will differ in finite sample cases. Their large sample behavior may not differ much and may be the same in some cases.

5. Asymptotic properties

The finite sample properties of R_δ^2 and R_k^2 depend on the values of b_δ and b_k. The moments of b_δ do not exist, see [2, p. 58] and [1]. The asymptotic distribution exists even when the exact distributions do not exist. So we derive the asymptotic distribution of the proposed statistics. We have considered a general structure of Σ_δ and Σ_ϕ and we also have assumed the existence and finiteness of the moments of ϕ and δ up to order four in terms of the respective coefficients of skewness and kurtosis. No form of the distributions of ϕ and δ are assumed. It is difficult to define the matrix variants of the measures of

skewness and kurtosis which are needed in deriving the asymptotic distributions of R_δ^2 and R_k^2. So for the sake of simplicity and for our purpose to derive the asymptotic distribution, we assume, without loss of generality, that all the data is mean corrected and $\Sigma_\delta = \sigma_\delta^2 I_p$ and $\Sigma_\phi = \sigma_\phi^2 I_n$. First we specify the distributional assumptions under this setup. We assume that the elements of Δ, δ_{ij}, $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, p)$ are independent and identically distributed random variables with mean 0, variance σ_δ^2, third moment $\gamma_{1\delta}\sigma_\delta^3$ and fourth moment $(\gamma_{2\delta}+3)\sigma_\delta^4$. Similarly, elements of Φ, ϕ_{ij}, $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, p)$ are assumed to be independent and identically distributed with first four finite moments given by 0, σ_ϕ^2, $\gamma_{1\phi}\sigma_\phi^3$ and $(\gamma_{2\phi}+3)\sigma_\phi^4$, respectively. Likewise it is also assumed that the elements of ϵ, ϵ_i, $(i = 1, 2, \ldots, n)$ are independent and identically distributed with first four finite moments given by 0, σ_ϵ^2, $\gamma_{1\epsilon}\sigma_\epsilon^3$ and $(\gamma_{2\epsilon}+3)\sigma_\epsilon^4$, respectively. Here, for a random variable Z, γ_{1Z} and γ_{2Z} denote the Pearson's coefficients of skewness and kurtosis of the random variable Z. Further, ϵ_i, δ_{ij} and ϕ_{ij} for all $i = 1, 2, \ldots, n, j = 1, 2, \ldots, p$ are also assumed to be statistically independent of each other. We further assume that the nth row of matrix M converges to σ_μ'. Consequently, we have $\lim_{n\to\infty} n^{-1}M'PM = \lim_{n\to\infty} n^{-1}M'M = \sigma_\mu\sigma_\mu' = \Sigma_\mu$. We define the following quantities which are useful in deriving the asymptotic distributions of R_δ^2 and R_k^2. Let

$$H = \sqrt{n}(S - \Sigma_x), \tag{5.1}$$

$$h = \sqrt{n}\{n^{-1}X'(\epsilon - \Delta\beta) + \sigma_\delta^2\beta\}, \tag{5.2}$$

$$g = \sqrt{n}\{n^{-1}(\epsilon - \Delta\beta)'(\epsilon - \Delta\beta) - \sigma_\epsilon^2 - \sigma_\delta^2\beta'\beta\}, \tag{5.3}$$

where $S = n^{-1}X'PX = n^{-1}X'X$. We present a lemma which is used to derive the asymptotic results:

Lemma 4. *Let $\{d_n\}$ be a sequence of $(p \times 1)$ non-stochastic vectors such that $\lim_{n\to\infty} d_n = d$. Then, as $n \to \infty$,*

$$\begin{pmatrix} Hd_n \\ h \\ g \end{pmatrix} \xrightarrow{d} N_{(2p+1)}\left(0, \begin{pmatrix} \Omega_H(dd') & \Omega_{hH}(d) & \Omega_{gH}(d) \\ \Omega_{hH}'(d) & \Omega_h & \Omega_{gh} \\ \Omega_{gH}'(d) & \Omega_{gh}' & \Omega_g \end{pmatrix}\right),$$

where

$$\begin{aligned} \Omega_H(dd') &= (\sigma_\delta^2 + \sigma_\phi^2)[\Sigma\{dd' + (d'd)I_p\} + dd'\sigma_\mu\sigma_\mu' + (d'\sigma_\mu\sigma_\mu'd)I_p] \\ &\quad + (\gamma_{1\phi}\sigma_\phi^3 + \gamma_{1\delta}\sigma_\delta^3)[f(\sigma_\mu\mathbf{1}_p', dd') + \{f(\sigma_\mu\mathbf{1}_p', dd')\}' \\ &\quad + 2f(I_p, \mathbf{1}_p\sigma_\mu dd')] + (\gamma_{2\phi}\sigma_\phi^4 + \gamma_{2\delta}\sigma_\delta^4)f(I_p, dd'), \end{aligned} \tag{5.4}$$

$$\begin{aligned} \Omega_{hH}(d) &= -\sigma_\delta^2[\Sigma(d\beta' + (d'\beta)I_p) + \gamma_{1\delta}\sigma_\delta\{f(\sigma_\mu\mathbf{1}_p', \beta d') \\ &\quad + f(I_p, \beta d'\sigma_\mu\mathbf{1}_p) + (f(\sigma_\mu\mathbf{1}_p', d\beta'))'\} + \gamma_{2\delta}\sigma_\delta^2 f(I_p, \beta d')], \end{aligned} \tag{5.5}$$

$$\Omega_h = (\sigma_\epsilon^2 + \sigma_\delta^2(\beta'\beta))\Sigma + \sigma_\delta^4\beta\beta' + \gamma_{1\delta}\sigma_\delta^3\{f(\sigma_\mu\mathbf{1}_p', \beta\beta') + \left(f(\sigma_\mu\mathbf{1}_p', \beta\beta')\right)'\} + \gamma_{2\delta}\sigma_\delta^4 f(I_p, \beta\beta'), \tag{5.6}$$

$$\Omega_g = 2(\sigma_\epsilon^2 + \sigma_\delta^2\beta'\beta)^2 + \gamma_{2\epsilon}\sigma_\epsilon^4 + \gamma_{2\delta}\sigma_\delta^4 f(I_p, \beta\beta')\beta, \tag{5.7}$$

$$\Omega_{gH}(d) = \left[\gamma_{1\delta}\sigma_\delta^3\{f(\sigma_\mu\mathbf{1}_p, d\beta') + f(I_p, \beta d'\sigma_\mu\mathbf{1}_p)\} + \gamma_{2\delta}\sigma_\delta^4 f(I_p, d\beta') + \sigma_\delta^2\beta d' + \sigma_\delta^4\beta' d\right]\beta \tag{5.8}$$

$$\Omega_{gh} = -2\sigma_\epsilon^2\sigma_\delta^2\beta - 2\sigma_\delta^4\beta\beta'\beta + \gamma_{1\epsilon}\sigma_\epsilon^3\sigma_\mu - \gamma_{1\delta}\sigma_\delta^3 f(\sigma_\mu\mathbf{1}_p', \beta\beta') - \gamma_{2\delta}\sigma_\delta^4 f(I_p, \beta\beta'), \tag{5.9}$$

where the function $f : \mathbb{R}^{p\times p} \times \mathbb{R}^{p\times p} \to \mathbb{R}^{p\times p}$ is defined as $f(Z_1, Z_2) = Z_1(Z_2 * I_p)$ for $Z_1, Z_2 \in \mathbb{R}^{p\times p}$, $*$ denotes the Hadamard product operator of matrices and $\xrightarrow{d}$ denotes the convergence in distribution.

Proof of Lemma 4 is detailed in the Appendix.

Theorem 2. *The asymptotic distribution of R_δ^2 as $n \longrightarrow \infty$ is given by*

$$\sqrt{n}\left(R_\delta^2 - \frac{\beta'\Sigma_x\beta}{\beta'\Sigma_x\beta + \sigma_\epsilon^2}\right) \xrightarrow{d} N\left(0, \frac{1}{(\beta'\Sigma\beta + \sigma_\epsilon^2)^4}\omega_\delta'\Omega_\delta\omega_\delta\right),$$

where $\Omega_\delta = \begin{pmatrix} \Omega_H(\beta\beta') & \Omega_{hH}(\beta) & \Omega_{gH}(\beta) \\ \Omega_{hH}'(\beta) & \Omega_h & \Omega_{gh} \\ \Omega_{gH}'(\beta) & \Omega_{gh}' & \Omega_g \end{pmatrix}$ and $\omega_\delta = \begin{pmatrix} \sigma_\epsilon^2\beta \\ 2\sigma_\epsilon^2(\Sigma - \sigma_\delta^2 I_p)^{-1}\beta \\ -\beta'\Sigma\beta \end{pmatrix}$. Here $\Omega_H(\beta\beta')$, $\Omega_{hH}(\beta)$ and $\Omega_{gH}(\beta)$ are obtained by replacing d by β in the covariance matrix given in Lemma 4.

Proof. Using (4.5), (2.1)–(2.4), (5.1)–(5.3), we have

$$\sqrt{n}\left(R_\delta^2 - \frac{\beta'\Sigma_x\beta}{\beta'\Sigma_x\beta+\sigma_\epsilon^2}\right) = \frac{1}{(\beta'\Sigma_x\beta+\sigma_\epsilon^2)^2}\left[\sigma_\epsilon^2\beta'H\beta + 2\sigma_\epsilon^2\beta'(\Sigma_x-\sigma_\delta^2 I_p)^{-1}h - \beta'\Sigma_x\beta g\right] + o_P(1)$$

$$= \frac{1}{(\beta'\Sigma_x\beta+\sigma_\epsilon^2)^2}\left(\sigma_\epsilon^2\beta', 2\sigma_\epsilon^2\beta'(\Sigma_x-\sigma_\delta^2 I_p)^{-1}, -\beta'\Sigma_x\beta\right)\begin{pmatrix}H\beta\\ h\\ g\end{pmatrix} + o_P(1). \tag{5.10}$$

Using (5.10), Slutsky's lemma and Lemma 4, we derive the asymptotic distribution of R_k^2 in the next theorem.

Theorem 3. *The asymptotic distribution of R_k^2 as $n \longrightarrow \infty$ is given by*

$$\sqrt{n}\left(R_k^2 - \frac{\beta'\Sigma_x\beta}{\beta'\Sigma_x\beta+\sigma_\epsilon^2}\right) \xrightarrow{d} N\left(0, \omega_k'\Omega_k\omega_k\right),$$

where

$$\omega_k = \begin{pmatrix}\sigma_\epsilon^2\beta\\ (\Sigma K_x)^{-1}\{\sigma_\epsilon^2 I_p - (I_p-K_x)\Sigma\beta'\beta\}\Sigma\beta\\ (\Sigma K_x)^{-1}\Sigma\beta(\sigma_\epsilon^2-\beta'\Sigma\beta)\\ \beta'\Sigma\beta\end{pmatrix},$$

$$\Omega_k = \begin{pmatrix}\Omega_H(\beta\beta') & \Omega_H(\beta\beta'(I_p-K_x)) & \Omega_{hH}(\beta) & \Omega_{gH}(\beta)\\ \Omega_H'(\beta\beta'(I_p-K_x)) & \Omega_H((I_p-K_x)\beta\beta'(I_p-K_x)) & \Omega_{hH}((I_p-K_x)\beta) & \Omega_{gH}((I_p-K_x)\beta)\\ \Omega_{hH}'(\beta) & \Omega_{hH}'((I_p-K_x)\beta) & \Omega_h & \Omega_{gh}\\ \Omega_{gH}'(\beta) & \Omega_{gH}'((I_p-K_x)\beta) & \Omega_{gh}' & \Omega_g\end{pmatrix},$$

and the elements of the asymptotic covariance matrix are obtained using (5.4)–(5.9) in Lemma 4.

Proof. Using (4.8), (2.1)–(2.4), (5.1)–(5.3), we have

$$\sqrt{n}\left(R_k^2 - \frac{\beta'\Sigma_x\beta}{\beta'\Sigma_x\beta+\sigma_\epsilon^2}\right) = \frac{1}{(\beta'\Sigma_x\beta+\sigma_\epsilon^2)^2}\left[\sigma_\epsilon^2\beta'H\beta + \beta'\Sigma_x\{\sigma_\epsilon^2 I_p - \beta'\beta\Sigma_x(I_p-K_x)\}(\Sigma_x K_x)^{-1}H(I_p-K_x)\beta\right.$$

$$\left. + (\sigma_\epsilon^2 - \beta'\Sigma_x\beta)\beta'\Sigma_x(\Sigma_x K_x)^{-1}h + \beta'\Sigma_x\beta g\right] + o_P(1)$$

$$= \frac{1}{(\beta'\Sigma_x\beta+\sigma_\epsilon^2)^2}\left(\sigma_\epsilon^2\beta', \beta'\Sigma_x\{\sigma_\epsilon^2 I_p - \beta'\beta\Sigma_x(I_p-K_x)\}(\Sigma_x K_x)^{-1},\right.$$

$$\left.(\sigma_\epsilon^2 - \beta'\Sigma_x\beta)\beta'\Sigma_x(\Sigma_x K_x)^{-1}, \beta'\Sigma_x\beta\right)\begin{pmatrix}H\beta\\ H(I_p-K_x)\beta\\ h\\ g\end{pmatrix} + o_P(1).$$

As a consequence of Lemma 4, it is easy to show that

$$\begin{pmatrix}H\beta\\ H(I_p-K_x)\beta\\ h\\ g\end{pmatrix} \xrightarrow{d} N_{(2p+1)}\left(0, \Omega_k\right).$$

Thus, using Slutsky's Lemma, the theorem is proved.

6. Simulation study

The asymptotic distributions of R_δ^2 and R_k^2 give us an idea about the behavior of their properties when the sample size grows large. In order to investigate the properties of these estimators in the finite samples, we conducted the Monte Carlo simulation experiments. To understand the effect of various involved variances more clearly, we assume that $\Sigma_\delta = \sigma_\delta^2 I_p$ and $\Sigma_\phi = \sigma_\phi^2 I_p$. We have also conducted the simulation for the general structures of Σ_δ and Σ_ϕ which can include the heteroskedastic and/or correlated errors and their findings are reported. Since we have not assumed any distributional assumptions like normality for any of the measurement errors or the random error component in deriving the asymptotic distributions, we consider the following three choices of the distributions based on their coefficients of skewness and kurtosis to conduct the simulations. We choose

(i) normal distribution,
(ii) t-distribution with 6 degrees of freedom and

Table 1
Absolute bias of R^2, R^2_δ, and R^2_k, when errors have normal distribution and $\theta = 0.9$.

	$(\sigma^2_\epsilon, \sigma^2_\phi, \sigma^2_\delta)$	AB(R^2)	AB(R^2_δ)	AB(R^2_k)
$n = 20$	(0.5, 0.5, 0.5)	0.2671	0.0451	0.0815
	(2.0, 0.5, 0.5)	0.2734	0.0450	0.0864
	(0.5, 2.0, 0.5)	0.1812	0.0250	0.0582
	(0.5, 0.5, 2.0)	0.6080	0.1953	0.1417
	(2.0, 2.0, 0.5)	0.1838	0.0255	0.0608
	(2.0, 0.5, 2.0)	0.6103	0.1922	0.1421
	(0.5, 2.0, 2.0)	0.4994	0.1525	0.1596
	(2.0, 2.0, 2.0)	0.4996	0.1496	0.1593
$n = 50$	(0.5, 0.5, 0.5)	0.3056	0.0244	0.0730
	(2.0, 0.5, 0.5)	0.3115	0.0258	0.0793
	(0.5, 2.0, 0.5)	0.1811	0.0151	0.0438
	(0.5, 0.5, 2.0)	0.6931	0.1061	0.1773
	(2.0, 2.0, 0.5)	0.1834	0.0156	0.0463
	(2.0, 0.5, 2.0)	0.6932	0.1037	0.1777
	(0.5, 2.0, 2.0)	0.5011	0.0470	0.1220
	(2.0, 2.0, 2.0)	0.5016	0.0475	0.1229

Table 2
Absolute bias of R^2, R^2_δ, and R^2_k, when errors have t distribution with 6 degrees of freedom and $\theta = 0.9$.

	$(\sigma^2_\epsilon, \sigma^2_\phi, \sigma^2_\delta)$	AB(R^2)	AB(R^2_δ)	AB(R^2_k)
$n = 20$	(0.5, 0.5, 0.5)	0.2812	0.0537	0.0943
	(2.0, 0.5, 0.5)	0.2842	0.0535	0.0969
	(0.5, 2.0, 0.5)	0.1920	0.0335	0.0691
	(0.5, 0.5, 2.0)	0.6108	0.2092	0.1428
	(2.0, 2.0, 0.5)	0.1933	0.0332	0.0703
	(2.0, 0.5, 2.0)	0.6121	0.2086	0.1445
	(0.5, 2.0, 2.0)	0.5049	0.1595	0.1647
	(2.0, 2.0, 2.0)	0.5045	0.1580	0.1643
$n = 50$	(0.5, 0.5, 0.5)	0.3180	0.0360	0.0860
	(2.0, 0.5, 0.5)	0.3217	0.0359	0.0901
	(0.5, 2.0, 0.5)	0.1894	0.0219	0.0524
	(0.5, 0.5, 2.0)	0.6917	0.1219	0.1750
	(2.0, 2.0, 0.5)	0.1901	0.0215	0.0531
	(2.0, 0.5, 2.0)	0.6924	0.1203	0.1765
	(0.5, 2.0, 2.0)	0.5133	0.0676	0.1367
	(2.0, 2.0, 2.0)	0.5136	0.0665	0.1374

(iii) gamma distribution with parameters (τ_1, τ_2, τ_3), where $\tau_1 = \tau_2\tau_3$ and $(\tau_2, \tau_3) \in \{(3, \sqrt{(\sigma^2_\epsilon/3)}), (3, \sqrt{(\sigma^2_\phi/3)}), (3, \sqrt{(\sigma^2_\delta/3)})\}$. Here τ_1 is the location parameter, τ_2 is the shape parameter and τ_3 is the scale parameter.

Note that the normal distribution has zero values for both the coefficients of skewness and kurtosis, gamma distribution has both nonzero values for the coefficients of skewness and kurtosis whereas t-distribution has only nonzero coefficient of kurtosis but zero coefficient of skewness. An inter comparison of the simulated values of bias and mean squared errors from these three distributions will give an idea about the effect of departure from normality on the properties of R^2_δ and R^2_k. Moreover, the random observations from the three distributions have been suitably scaled so that for any particular combination of the variances, the generated random values from all the three distributions have the same means and the same variances. Various combinations of the values of variances (σ^2_ϵ, σ^2_ϕ, and σ^2_δ) are considered so as to reflect the broad spectrum of smaller variances to large variances and also a combination of them. We obtained the empirical absolute bias and the empirical mean squared errors of R^2, R^2_δ and R^2_k based on 25,000 repetitions. Since our main interest is in the magnitude of bias, we have considered the absolute bias. The empirical values of absolute bias are presented in Tables 1–3 and the empirical mean squared errors are presented in Tables 4–6 based on the chosen sample sizes $n = 20, 50$. In order to save space, we are not providing the matrix M here. However, as per the assumption taken in Section 5, we adopted such matrix M as the nth row of it which converges to $\sigma'_\mu = \begin{pmatrix} 0 & 2 & 4 & 1 & 3 \end{pmatrix}$, as $n \to \infty$.

The statistics R^2, R^2_δ and R^2_k are expected to reflect the goodness of fit. If the model itself has certain defects and measurement errors are also too high, then a lower value of R^2, R^2_δ and R^2_k may not truly reveal the goodness of fit as the lower values may be occurring due to in built model defects as well. In order to avoid such situations, we have first considered a good model to generate the data with higher value of θ so that the values of R^2, R^2_δ and R^2_k reflect the effect of measurement errors only. This is done in the results presented in Tables 1–7 where $\theta = 0.9$ is considered for generating the data. Next, we also studied the effect of measurement errors on the values of R^2, R^2_δ and R^2_k when the data is generated from a reasonably large range of values of θ. This is done for the structural model and the results are presented in Table 8.

Table 3
Absolute bias of R^2, R^2_δ, and R^2_k, when errors have gamma distribution and $\theta = 0.9$.

	$(\sigma^2_\epsilon, \sigma^2_\phi, \sigma^2_\delta)$	AB(R^2)	AB(R^2_δ)	AB(R^2_k)
$n = 20$	(0.5, 0.5, 0.5)	0.2777	0.0514	0.0909
	(2.0, 0.5, 0.5)	0.2814	0.0515	0.0939
	(0.5, 2.0, 0.5)	0.1899	0.0314	0.0669
	(0.5, 0.5, 2.0)	0.6092	0.2079	0.1427
	(2.0, 2.0, 0.5)	0.1908	0.0311	0.0677
	(2.0, 0.5, 2.0)	0.6112	0.2062	0.1425
	(0.5, 2.0, 2.0)	0.5049	0.1559	0.1644
	(2.0, 2.0, 2.0)	0.5026	0.1551	0.1617
$n = 50$	(0.5, 0.5, 0.5)	0.3136	0.0317	0.0815
	(2.0, 0.5, 0.5)	0.3182	0.0327	0.0863
	(0.5, 2.0, 0.5)	0.1871	0.0197	0.0499
	(0.5, 0.5, 2.0)	0.6907	0.1180	0.1730
	(2.0, 2.0, 0.5)	0.1885	0.0195	0.0514
	(2.0, 0.5, 2.0)	0.6920	0.1175	0.1762
	(0.5, 2.0, 2.0)	0.5094	0.0614	0.1319
	(2.0, 2.0, 2.0)	0.5114	0.0618	0.1346

Table 4
Mean squared error of R^2, R^2_δ, and R^2_k, when errors have normal distribution and $\theta = 0.9$.

	$(\sigma^2_\epsilon, \sigma^2_\phi, \sigma^2_\delta)$	MSE(R^2)	MSE(R^2_δ)	MSE(R^2_k)
$n = 20$	(0.5, 0.5, 0.5)	0.0751	0.0043	0.0107
	(2.0, 0.5, 0.5)	0.0788	0.0043	0.0120
	(0.5, 2.0, 0.5)	0.0351	0.0012	0.0058
	(0.5, 0.5, 2.0)	0.3801	0.0636	0.0353
	(2.0, 2.0, 0.5)	0.0363	0.0013	0.0064
	(2.0, 0.5, 2.0)	0.3830	0.0622	0.0354
	(0.5, 2.0, 2.0)	0.2612	0.0435	0.0429
	(2.0, 2.0, 2.0)	0.2613	0.0422	0.0428
$n = 50$	(0.5, 0.5, 0.5)	0.0954	0.0011	0.0076
	(2.0, 0.5, 0.5)	0.0993	0.0013	0.0089
	(0.5, 2.0, 0.5)	0.0338	0.0004	0.0029
	(0.5, 0.5, 2.0)	0.4862	0.0219	0.0459
	(2.0, 2.0, 0.5)	0.0348	0.0005	0.0033
	(2.0, 0.5, 2.0)	0.4863	0.0210	0.0458
	(0.5, 2.0, 2.0)	0.2560	0.0043	0.0220
	(2.0, 2.0, 2.0)	0.2566	0.0045	0.0223

Table 5
Mean squared error of R^2, R^2_δ, and R^2_k, when errors have t distribution with 6 degrees of freedom and $\theta = 0.9$.

	$(\sigma^2_\epsilon, \sigma^2_\phi, \sigma^2_\delta)$	MSE(R^2)	MSE(R^2_δ)	MSE(R^2_k)
$n = 20$	(0.5, 0.5, 0.5)	0.0840	0.0058	0.0142
	(2.0, 0.5, 0.5)	0.0859	0.0059	0.0150
	(0.5, 2.0, 0.5)	0.0401	0.0023	0.0082
	(0.5, 0.5, 2.0)	0.3838	0.0713	0.0356
	(2.0, 2.0, 0.5)	0.0408	0.0024	0.0085
	(2.0, 0.5, 2.0)	0.3857	0.0712	0.0368
	(0.5, 2.0, 2.0)	0.2672	0.0464	0.0455
	(2.0, 2.0, 2.0)	0.2668	0.0457	0.0456
$n = 50$	(0.5, 0.5, 0.5)	0.1039	0.0026	0.0106
	(2.0, 0.5, 0.5)	0.1064	0.0026	0.0115
	(0.5, 2.0, 0.5)	0.0373	0.0010	0.0043
	(0.5, 0.5, 2.0)	0.4844	0.0272	0.0450
	(2.0, 2.0, 0.5)	0.0377	0.0010	0.0044
	(2.0, 0.5, 2.0)	0.4853	0.0264	0.0457
	(0.5, 2.0, 2.0)	0.2692	0.0087	0.0274
	(2.0, 2.0, 2.0)	0.2697	0.0086	0.0279

First we analyze the values of absolute bias and mean squared errors of R^2 from Tables 1–3. Looking at these values, we observe that the values of the absolute bias and the mean squared errors of R^2 do not change when the sample size increases and it remains true for all the considered combinations of variances $(\sigma^2_\epsilon, \sigma^2_\phi, \sigma^2_\delta)$ and under all the three distributions. On

Table 6
Mean squared error of R^2, R^2_δ, and R^2_k, when errors have gamma distribution and $\theta = 0.9$.

	$(\sigma^2_\epsilon, \sigma^2_\phi, \sigma^2_\delta)$	MSE(R^2)	MSE(R^2_δ)	MSE(R^2_k)
$n = 20$	(0.5, 0.5, 0.5)	0.0816	0.0053	0.0131
	(2.0, 0.5, 0.5)	0.0838	0.0053	0.0140
	(0.5, 2.0, 0.5)	0.0390	0.0020	0.0076
	(0.5, 0.5, 2.0)	0.3818	0.0705	0.0358
	(2.0, 2.0, 0.5)	0.0396	0.0020	0.0079
	(2.0, 0.5, 2.0)	0.3841	0.0698	0.0358
	(0.5, 2.0, 2.0)	0.2669	0.0442	0.0449
	(2.0, 2.0, 2.0)	0.2646	0.0446	0.0442
$n = 50$	(0.5, 0.5, 0.5)	0.1009	0.0019	0.0095
	(2.0, 0.5, 0.5)	0.1039	0.0021	0.0105
	(0.5, 2.0, 0.5)	0.0363	0.0007	0.0038
	(0.5, 0.5, 2.0)	0.4828	0.0254	0.0439
	(2.0, 2.0, 0.5)	0.0369	0.0008	0.0041
	(2.0, 0.5, 2.0)	0.4848	0.0258	0.0457
	(0.5, 2.0, 2.0)	0.2650	0.0071	0.0256
	(2.0, 2.0, 2.0)	0.2671	0.0074	0.0265

Table 7
Absolute bias and mean squared error when random terms have multivariate normal distribution and $\theta = 0.9$.

n	AB(R^2)	AB(R^2_δ)	AB(R^2_k)	MSE(R^2)	MSE(R^2_δ)	MSE(R^2_k)
20	0.4030	0.2458	0.1410	0.1741	0.0996	0.0339
50	0.4764	0.1932	0.1717	0.2337	0.0620	0.0390

Table 8
Absolute bias and mean squared error for various values of θ.

θ	n	AB(R^2)	AB(R^2_δ)	AB(R^2_k)	MSE(R^2)	MSE(R^2_δ)	MSE(R^2_k)
0.1265	20	0.0266	0.5521	0.7003	0.0042	0.3588	0.5143
0.1265	250	0.1037	0.4101	0.1424	0.0110	0.2426	0.0371
0.4200	20	0.2643	0.2693	0.4141	0.0732	0.1264	0.1936
0.4200	250	0.3876	0.1836	0.0696	0.1506	0.1000	0.0256
0.7655	20	0.6557	0.1232	0.1262	0.4314	0.0778	0.0350
0.7655	250	0.7185	0.0569	0.1054	0.5168	0.0328	0.0294
0.9988	20	0.8812	0.3401	0.0858	0.7779	0.1743	0.0171
0.9988	250	0.8632	0.0530	0.0408	0.7456	0.0058	0.0032

the other hand, the values of the absolute bias and the values of the mean squared errors of R^2_δ and R^2_k decrease when the sample size increases. This verifies the analytical results proved in the earlier section. This is due to the reason that R^2 is an inconsistent estimator of θ, while both R^2_δ and R^2_k are the consistent estimators of θ. Moreover, the values of absolute bias and mean squared errors of R^2_δ and R^2_k are much smaller than the values of R^2. In fact, the obtained values of R^2 are not correct and if considered, they may lead to incorrect statistical conclusions. This clearly indicates that the values of R^2 are not suitable at all for judging the goodness of fit in the measurement error models.

Next we analyze the values of absolute bias of R^2, R^2_δ, and R^2_k from Tables 1–3 under different distributions. First we consider the results from the normally distributed errors. We notice that the values of absolute bias of R^2 are much higher than the values of absolute bias of R^2_δ and R^2_k. Since the values of R^2 are based on an inconsistent estimator, they are not suitable and that is why we study the behavior of absolute bias of R^2_δ and R^2_k only. We observe that when only the value of σ^2_ϵ increases and the values of other two variances remain the same, the values of absolute bias of R^2_δ and R^2_k do not differ much and they decrease slightly when the value of σ^2_ϕ increases while the other two variances remain fixed. When the values of σ^2_δ increase keeping the values of the other two variances fixed, then the values of absolute bias of R^2_δ and R^2_k increase and both the values are affected severely. The impact of the values of σ^2_δ on the absolute bias of R^2_δ and R^2_k is more than the values of σ^2_ϵ or σ^2_ϕ. When both the values of σ^2_ϵ and σ^2_ϕ increase together and the value of σ^2_δ is kept fixed, then the values of absolute bias of R^2_δ and R^2_k decrease. On the other hand, when the values of σ^2_ϵ and σ^2_δ increase together while the value of σ^2_ϕ stays fixed, then the values of absolute bias of R^2_δ and R^2_k increase and are severely affected. In case, if only one of the values out of σ^2_ϕ and σ^2_δ increases, then the absolute bias again increases but it is clear that the values of σ^2_δ affect it more and tend to increase the magnitude of bias. The trends and effects of the values of σ^2_ϵ, σ^2_δ and σ^2_ϕ on the values of absolute bias in case of t-distributed errors and gamma distributed errors are similar to the case of normally distributed errors. When comparing

the values under normal distributions with t- and gamma distributed errors, we find that there is significant difference in the corresponding values. On the other hand, there is a significant difference in the values of absolute bias under the normal and the t-distributed errors but not much difference is present between the values under the t-distributed errors and the values under the gamma distributed errors. This clearly indicates that the departure from normality does affect the performance of R^2_δ and R^2_k. Moreover, the skewness of the distribution has more impact on the magnitude of bias than the kurtosis of the distribution. When the sample size increases, the magnitude of respective bias decreases.

Now we analyze the behavior of the values of empirical mean squared errors from Tables 4–6. Again, since R^2 is an inconsistent estimator of θ and its values are not reliable, we concentrate only on the values of empirical mean squared errors of R^2_δ and R^2_k. We notice that when the values of σ^2_ϵ increase alone keeping the values of the other two variances fixed, then the mean squared errors of R^2_δ and R^2_k do not change significantly. The values of the mean squared errors of R^2_δ and R^2_k decrease when the values of σ^2_ϕ increase. On the other hand, the values of mean squared errors increase and are severely affected when the values of σ^2_δ increase. When both the values of σ^2_ϵ and σ^2_δ increase together and the values of σ^2_ϕ are kept fixed, then the values of mean squared errors of R^2_δ and R^2_k decrease. In other cases, when both the values of σ^2_ϵ and σ^2_ϕ increase while the values of σ^2_δ stay fixed or both the values of σ^2_δ and σ^2_ϕ increase while the values of σ^2_ϵ stay fixed, then the mean squared errors of both R^2_δ as well as R^2_k increase. The values of mean squared errors of R^2_δ and R^2_k increase more with the increase in the values of σ^2_δ than the values of the other two variances. When the values of all the variances increase simultaneously, then also the mean squared errors of R^2_δ and R^2_k increase. When the sample size increases, then the values of mean squared errors of both R^2_δ and R^2_k decrease in all the cases. The trends and effects of σ^2_ϵ, σ^2_δ and σ^2_ϕ in the cases of t- and gamma distributed errors are the same as in the case of normally distributed errors. Comparing the respective values of the mean squared errors under different distributions, we notice that the values of mean squared errors under the normal distribution are smaller than the corresponding values under the t- and gamma distributed errors. Moreover, the values of mean squared errors under the t- and gamma distributed errors do not differ much. This clearly indicates that the departure from normality affects the mean squared errors of R^2_δ and R^2_k both. Moreover, the departure from the symmetry influences the mean squared errors more than the departure from the peakedness of the distributions.

Further, in order to analyze the performance of R^2_δ and R^2_k in the finite samples when the matrices $\Sigma_\delta \neq \sigma^2_\delta I$ and $\Sigma_\phi \neq \sigma^2_\phi I$, i.e., they are not in the form of identity matrices, we conducted some more simulations. For illustration, we are reporting here one set of results which is conducted using the normal distribution of all the errors ϵ, ϕ and δ with $\sigma^2_\epsilon = 2$,

$$\Sigma_\delta = \begin{pmatrix} 1.00 & 0.09 & -0.03 & 0.01 & -0.03 \\ 0.09 & 0.98 & 0.00 & -0.08 & -0.03 \\ -0.03 & 0.00 & 0.96 & 0.01 & -0.03 \\ 0.01 & -0.08 & 0.01 & 1.09 & 0.04 \\ -0.03 & -0.03 & -0.03 & 0.04 & 1.05 \end{pmatrix} \quad \text{and}$$

$$\Sigma_\phi = \begin{pmatrix} 1.15 & -0.08 & 0.04 & -0.03 & 0.02 \\ -0.08 & 1.08 & 0.06 & 0.00 & 0.00 \\ 0.04 & 0.06 & 0.93 & 0.02 & 0.08 \\ -0.01 & 0.00 & 0.02 & 1.02 & -0.03 \\ -0.02 & 0.00 & 0.08 & -0.03 & 0.91 \end{pmatrix}.$$

The results of the simulation outcomes on the empirical absolute bias and the empirical mean squared errors are presented in Table 7. We observe that the results are more or less similar to the cases when $\Sigma_\delta = \sigma^2_\delta I$ and $\Sigma_\phi = \sigma^2_\phi I$ and they have the similar pattern and the similar conclusions except that the values of absolute bias and mean squared error of R^2, R^2_δ and R^2_k differ. It is clear that our suggested forms of R^2_δ and R^2_k work well with any form of Σ_δ and Σ_ϕ in measurement error models.

The results in Table 8 present the values of absolute bias and mean squared errors of R^2, R^2_δ and R^2_k for reasonably lower value of θ $(= 0.1265)$ to reasonably higher value of θ $(= 0.9988)$ for sample sizes 20 and 250 under the setup of the structural model. It is clear that the values of R^2 are bad enough to be considered as its absolute bias and mean squared error increase as the sample size increases which is because it is an inconsistent estimator of θ. So we consider the values for R^2_δ and R^2_k. Their absolute bias and mean squared error decrease as the sample size increases for all values of θ. It is clear from the results that the proposed statistics R^2_δ and R^2_k work well for all the values of θ. The values of absolute bias and mean squared error of R^2_δ and R^2_k decrease as θ increases. An interesting observation emerges as follows. When θ is low, then the absolute bias of R^2_δ is more than that of R^2_k. When θ is high, then the reverse holds true, i.e., absolute bias of R^2_δ is higher than that of R^2_k. When θ is low, then the mean squared error of R^2_k is more than that of R^2_δ. On the other hand, when θ is high, then the mean squared error of R^2_k is smaller than that of R^2_δ. Such information can be used in creating the guidelines for practitioners. Suppose a practitioner wants to know that which of the information between the known covariance matrix or known reliability matrix is to be used for fitting the model, then if the practitioner has some prior information about the value of θ, then he can decide as follows: If θ is expected to be low, then use R^2_δ and if θ is expected to be high, then use R^2_k. It may be noted that for low values of θ, quite large values of n are required before R^2 performs worse than the other two

consistent coefficients of determination, viz., R_δ^2 and R_k^2. It may be noted that quite large values of n may not even help to perform R^2 better than the other two consistent coefficients of determination, viz., R_δ^2 and R_k^2. The reason being that the R^2 is an inconsistent estimator of θ and it will never converge to θ howsoever large n may be.

7. Conclusion

The conventional coefficient of determination becomes inconsistent for the population version of the coefficient of determination given by θ in the presence of measurement errors in the data. So we have proposed two forms of goodness of fit statistics which can be used to judge the goodness of fit in measurement error models. These statistics are based on the utilization of two variants of additional information and the conventional form of the coefficient of determination. The additional information is assumed to be available from outside the sample in the form of the known covariance matrix of measurement errors in the explanatory variables and the known reliability matrix associated with the explanatory variables. We also have established a connection between the two cases, i.e., when the reliability matrix is estimated from the data and when the covariance matrix of the measurement error associated with the explanatory variables is known. Thus obtained statistics like coefficients of determination are consistent for estimating θ and can be used to judge the goodness of fit in measurement error models. Due to the issues like the first moment of b_δ does not exist (see [1]) it is difficult to define the coefficient of determination for measurement error models as it is done through analysis of variance in the linear regression analysis in no-measurement error situations. The asymptotic distributions of R_δ^2 and R_k^2 are derived. The finite sample properties R_δ^2 and R_k^2 are studied through the Monte-Carlo simulation experiments under the situations when $\Sigma_\delta = \sigma_\delta^2 I$, $\Sigma_\phi = \sigma_\phi^2 I$ and for any general structure of Σ_δ and Σ_ϕ. It is observed that the values of empirical bias and empirical mean squared errors of R_δ^2 and R_k^2 turn out to be satisfactory even for a sample of size 20. The simulated results for the cases when $\Sigma_\delta \neq \sigma_\delta^2 I$ and $\Sigma_\phi \neq \sigma_\phi^2 I$ are also satisfactory. The values of R_δ^2 and R_k^2 also work well for all the values of θ. Moreover, R_δ^2 is preferable for the lower values of θ and R_k^2 is preferable for the higher values of θ. So the proposed R_δ^2 and R_k^2 can be used in real data modeling to judge the goodness of fit in the measurement error models for all the values of θ. It is to be noticed that the values of absolute bias and mean squared errors of R_δ^2 and R_k^2 are affected by the departure from normality. So one has to be cautious in using the proposed statistics when the validity of the normal distribution assumption for errors is doubtful.

In case, the values of Σ_δ or K_x are not available, then one possible solution is to estimate if the repeated observations are available. The forms of R_δ^2 and R_k^2 can be obtained by replacing Σ_δ or K_x by their respective consistent estimates. The resulting statistics will remain consistent. The investigation of their properties is out of the purview of this paper.

Appendix

Lemma 5 (*Central Limit Theorem*). *Let $V_n = \sum_{i=1}^n D_{in}\psi_i$, where $\psi_1, \psi_2, \ldots, \psi_n$ are $p \times 1$ independent and identically distributed random vectors with $E(\psi_i) = 0$ and $D_{1n}, D_{2n}, \ldots, D_{nn}$ are $q \times p$ non-stochastic matrices. Suppose that $\lim_{n\to\infty} \mathrm{cov}(V_n) = \Omega$, where $|(\Omega)_{ij}| < \infty$ and Ω is positive definite. If there exists a function $\varphi(n)$ such that $\lim_{n\to\infty}\varphi(n) = \infty$ and if the elements of $\varphi(n)D_{in}$ are bounded, then V_n has a limiting q-variate normal distribution with mean vector 0 and covariance matrix Ω, see* [20, *p. 212*].

Proof of Lemma 4. Using vec operator and Kronecker's product of matrices, we can write

$$Hd_n = \sum_{i=1}^n A_{in}u_i, \tag{7.1}$$

where, for $i = 1, 2, \ldots, n$, $A_{in} = n^{-1/2}(d_n' \otimes I_p)[(I_p \otimes \mu_i), (\mu_i' \otimes I_{p^2}), I_{p^2}, I_{p^2}, I_{p^2}]$ are $p \times (p + 3p^2 + p^3)$ non-stochastic matrices and

$$u_i = \begin{pmatrix} (\phi_i + \delta_i) \\ \mathrm{vec}(I_p \otimes (\phi_i + \delta_i)) \\ \mathrm{vec}(\phi_i\delta_i' + \delta_i\phi_i') \\ \mathrm{vec}(\phi_i\phi' - \sigma_\phi^2 I_p) \\ \mathrm{vec}(\delta_i\delta_i' - \sigma_\delta^2 I_p) \end{pmatrix}$$

are $((p + 3p^2 + p^3) \times 1)$ independently and identically distributed random vectors.

Similarly, we can write

$$h = \sum_{i=1}^n B_{in}v_i, \tag{7.2}$$

where, for $i = 1, 2, \ldots, n$, $B_{in} = n^{-1/2}[\mu_i, I_p, I_p, -(\beta' \otimes I_p)(I_p \otimes \mu_i), -(\beta' \otimes I_p), -(\beta' \otimes I_p)]$ are $p \times (1 + 3p + 2p^2)$ non-stochastic matrices and

$$v_i = \begin{pmatrix} \epsilon_i \\ \delta_i\epsilon_i \\ \phi_i\epsilon_i \\ \delta_i \\ \mathrm{vec}(\phi_i\delta_i') \\ \mathrm{vec}(\delta_i\delta_i' - \sigma_\delta^2 I_p) \end{pmatrix}$$

are $((1 + 3p + 2p^2) \times 1)$ independently and identically distributed random vectors.

In the similar manner we write

$$g = \sum_{i=1}^{n} C_{in} w_i, \tag{7.3}$$

where, for $i = 1, 2, \ldots, n$, $C_{in} = n^{-1/2}[1, 2\beta', \beta'(\beta' \otimes I_p)]$ are $(1 \times (1 + p + p^2))$ non-stochastic vectors and

$$w_i = \begin{pmatrix} \epsilon_i^2 - \sigma_\epsilon^2 \\ \delta_i\epsilon_i \\ \mathrm{vec}(\delta_i\delta_i' - \sigma_\delta^2 I_p) \end{pmatrix}$$

are $((1 + p + p^2) \times 1)$ independently and identically distributed random vectors.

Combining (7.1)–(7.3), we have

$$\begin{pmatrix} Hd_n \\ h \\ g \end{pmatrix} = \sum_{i=1}^{n} \begin{pmatrix} A_{in} & 0 & 0 \\ 0 & B_{in} & 0 \\ 0 & 0 & C_{in} \end{pmatrix} \begin{pmatrix} u_i \\ v_i \\ w_i \end{pmatrix}, \tag{7.4}$$

where, for $i = 1, 2, \ldots, n$,

$$\begin{pmatrix} A_{in} & 0 & 0 \\ 0 & B_{in} & 0 \\ 0 & 0 & C_{in} \end{pmatrix}$$

is the $((2p + 1) \times (2 + 5p + 6p^2 + 3p^3))$ non-stochastic matrix and

$$\begin{pmatrix} u_i \\ v_i \\ w_i \end{pmatrix}$$

are $((2+5p+6p^2+3p^3) \times 1)$ independently and identically distributed random vectors. Using the assumption about matrix M, it can be seen that the element of matrix

$$\sqrt{n}\begin{pmatrix} A_{in} & 0 & 0 \\ 0 & B_{in} & 0 \\ 0 & 0 & C_{in} \end{pmatrix}$$

is bounded. Using the assumptions about the moments of random terms in the model, we have

$$E\begin{pmatrix} Hd_n \\ h \\ g \end{pmatrix} = 0,$$

and $\lim_{n\to\infty} E(Hd_nd_n'H) = \Omega_H(dd')$, $\lim_{n\to\infty} E(hh') = \Omega_h$, $\lim_{n\to\infty} E(g^2) = \Omega_g$, $\lim_{n\to\infty} E(hd_n'H) = \Omega_{hH}(d)$, $\lim_{n\to\infty} E(Hd_ng) = \Omega_{gH}(d)$, $\lim_{n\to\infty} E(hg) = \Omega_{gh}$. Therefore, on using the central limit theorem given in Lemma 5 for the function $\varphi(n) = \sqrt{n}$, we conclude that $\begin{pmatrix} Hd_n \\ h \\ g \end{pmatrix}$ has a $(2p + 1)$-variate limiting normal distribution with mean vector 0 and covariance matrix $\begin{pmatrix} \Omega_H(d) & \Omega_{hH}(d) & \Omega_{gH}(d) \\ \Omega_{hH}'(d) & \Omega_h & \Omega_{gh} \\ \Omega_{gH}'(d) & \Omega_{gh}' & \Omega_g \end{pmatrix}$.

References

[1] C.-L. Cheng, A. Kukush, Non-existence of the first moment of the adjusted least squares estimator in multivariate errors-in-variables model, Metrika 64 (1) (2006) 41–46.
[2] C.-L. Cheng, J.W. Van Ness, Statistical Regression with Measurement Errors, Arnold, London, 1999.

[3] C.L. Cheng, J.W. Van Ness, On the unreplicated ultrastructural model, Biometrika 78 (1991) 442–445.
[4] J.S. Cramer, Mean and variance of R^2 in small and moderate samples, J. Econometrics 35 (1987) 253–266.
[5] G.R. Dolby, The ultrastructural relation: a synthesis of the functional and structural relations, Biometrika 63 (1976) 39–50.
[6] N. Eshima, M. Tabata, Entropy coefficient of determination for generalized linear models, Comput. Statist. Data Anal. 54 (5) (2010) 1381–1389.
[7] N. Eshima, M. Tabata, Three predictive power measures for generalized linear models: the entropy coefficient of determination, the entropy correlation coefficient and the regression correlation coefficient, Comput. Statist. Data Anal. 55 (11) (2011) 3049–3058.
[8] W.A. Fuller, Measurement Error Models, John Willey, 1987.
[9] L.J. Gleser, The importance of assessing measurement reliability in multivariate regression, J. Amer. Statist. Assoc. 87 (419) (1992) 696–707.
[10] L.J. Gleser, Estimators of slopes in linear errors-in-variables regression models when the predictors have known reliability matrix, Statist. Probab. Lett. 17 (1993) 113–121.
[11] L.J. Gleser, A note on G.R. Dolby's ultrastructural model, Biometrika 72 (1985) 117–124.
[12] G.J. Hahn, The coefficient of determination exposed!, Chem. Technol. 3 (1973) 609–614.
[13] J.E. Hilliard, W.P. Lloyd, Coefficient of determination in a simultaneous equation model: a pedagogic note, J. Bus. Res. 8 (1) (1980) 1–6.
[14] C.S. Hong, J.H. Ham, H.I. Kim, Variable selection for logistic regression model using adjusted coefficients of determination, Korean J. Appl. Statist. 18 (2) (2005) 435–443.
[15] O. Hössjer, On the coefficient of determination for mixed regression models, J. Statist. Plann. Inference 138 (10) (2008) 3022–3038.
[16] L.-S. Huang, J. Chen, Analysis of variance, coefficient of determination and F-test for local polynomial regression, Ann. Statist. 36 (5) (2008) 2085–2109.
[17] J.L. Knight, The coefficient of determination and simultaneous equation systems, J. Econometrics 14 (2) (1980) 265–270.
[18] J.G. Liao, D. McGee, Adjusted coefficients of determination for logistic regression, American Statist. 57 (3) (2003) 161–165.
[19] S.R. Lipsitz, T. Leong, J. Ibrahim, S. Lipshultz, A partial correlation coefficient and coefficient of determination for multivariate normal repeated measures data, Statistician 50 (1) (2001) 87–95.
[20] E. Malinvaud, Statistical Methods of Econometrics, North-Holland Publishing Co, Amsterdam, 1966.
[21] É. Marchand, Point estimation of the coefficient of determination, Statist. Decis. 19 (2) (2001) 137–154.
[22] É Marchand, On moments of beta mixtures, the noncentral beta distribution, and the coefficient of determination, J. Stat. Comput. Simul. 59 (2) (1997) 161–178.
[23] J.W. McKean, G.L. Sievers, Coefficients of determination for least absolute deviation analysis, Statist. Probab. Lett. 5 (1) (1987) 49–54.
[24] N.J.D. Nagelkerke, A note on a general definition of the coefficient of determination, Biometrika 78 (3) (1991) 691–692.
[25] K. Ohtani, The density functions of R^2 and $\bar{R}^2$, and their risk performance under asymmetric loss in misspecified linear regression models, Econom. Model. 11 (4) (1994) 463–471.
[26] K. Ohtani, D.E.A. Giles, The absolute error risks of regression "goodness of fit measures", J. Quant. Econom. 12 (1996) 17–26.
[27] K. Ohtani, H. Hasegawa, On small scale properties of R^2 in a linear regression model with multivariate t errors and proxy variables, Econom. Theory 9 (1993) 504–515.
[28] O. Renaud, M.-P. Victoria-Feser, A robust coefficient of determination for regression, J. Statist. Plann. Inference 140 (7) (2010) 1852–1862.
[29] H. Schneeweiss, Consistent estimation of a regression with errors in the variables, Metrika 23 (1976) 101–115.
[30] H. Schneeweiss, Note on a linear model with errors in the variables and with trend, Statist. Papers 32 (1991) 261–264.
[31] Shalabh, Improved estimation in measurement error models through Stein-rule procedure, J. Multivariate Anal. 67 (1998) 35–48.
[32] Shalabh, Corrigendum, J. Multivariate Anal. 74 (2000) 162.
[33] Shalabh, Consistent estimation of coefficients in measurement error models under non-normality, J. Multivariate Anal. 86 (2) (2003) 227–241.
[34] Shalabh, G. Garg, N. Misra, Restricted regression estimation in measurement error models, Computational Statistics and Data Analysis 52 (2) (2007) 1149–1166.
[35] Shalabh, G. Garg, N. Misra, Use of prior information in the consistent estimation of regression coefficients in measurement error models, J. Multivariate Anal. 100 (7) (2009) 1498–1520.
[36] Shalabh, G. Garg, N. Misra, Consistent estimation of regression coefficients in measurement error model using stochastic a priori information, Statist. Papers 51 (2010) 717–748.
[37] Shalabh, G. Garg, N. Misra, Estimation of regression coefficients in a restricted measurement error model using instrumental variables, Commun. Stat.: - Theory Methods 40 (2011) 3614–3629.
[38] M.D Smith, Comparing approximations to the expectation of quadratic forms in normal variables, Econometric Rev. 15 (1996) 81–95.
[39] A.K. Srivastava, Shobhit, A family of estimators for the coefficient of determination in linear regression models, J. Appl. Statist. Sci. 11 (2) (2002) 133–144.
[40] A.K. Srivastava, V.K. Srivastava, A. Ullah, The coefficient of determination and its adjusted version in linear regression models, Econometric Reviews 14 (1995) 229–240.
[41] J.S. Tanaka, G.J. Huba, A general coefficient of determination for covariance structure models under arbitrary GLS estimation, British J. Math. Statist. Psych. 42 (2) (1989) 233–239.
[42] T. Tjur, Coefficients of determination in logistic regression models — a new proposal: the coefficient of discrimination, American Statistician 63 (4) (2009) 366–372.
[43] A. Ullah, V.K. Srivastava, Moments of the ratio of quadratic forms in non-normal variables with econometric examples, Journal of Econometrics 62 (1994) 129–142.
[44] A. van der Linde, G. Tutz, On association in regression: the coefficient of determination revisited, Statistics 42 (1) (2008) 1–24.