Coefficient of Determination for Multiple Measurement Error Models

Shalabh[1,2] and G. Garg[3]
[1]Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur - 208 016, INDIA,
[3]Decision Sciences Area, Indian Institute of Management Lucknow, Lucknow - 226 013, INDIA,
[2]Corresponding author: Shalabh, e-Mail: shalab@iitk.ac.in, shalabh1@yahoo.com

Abstract

The coefficient of determination (R^2) is used for judging the goodness of fit in a linear regression model. It gives valid results only when the observations are correctly observed without any measurement error. The R^2 provides invalid results in the presence of measurement errors in the data in the sense that sample R^2 becomes an inconsistent estimator of population multiple correlation coefficient between the study variable and explanatory variables. The corresponding variants of R^2 which can be used to judge the goodness of fit in multivariate measurement error model have been proposed in this paper. These variants are based on the utilization of information on known covariance matrix of measurement errors and known reliability matrix associated with explanatory variables. The asymptotic properties of the traditional R^2 and proposed R^2 have been studied.
Key Words: Linear regression, ultrastructural model, non-normal distribution.

1 Introduction

A popular tool to determine the adequacy of the linear regression model is the coefficient of determination (R^2) which is based on the proportion of variability of the study variable that can be explained through the knowledge of a given set of variables. R^2 is essentially the multiple correlation coefficient between the study variable and all the explanatory variables present in the linear regression model. Therefore, R^2 reflects the capability of the regression relationship to predict the values of study variable.

When measurement error creeps into the data, the optimal properties of the estimators in a linear regression model are disturbed. The presence of measurement error turns the ordinary least squares estimator (OLSE) of regression coefficient which is the best linear unbiased estimator into a biased and inconsistent estimator. The properties of coefficient of determination, R^2, are also disturbed. The value of R^2 obtained by ignoring the measurement errors may lead to incorrect statistical inference. So a question arises that how to judge the goodness of fit in a linear regression model when the observations are contaminated with measurement errors.

Since the form of R^2 is directly related to estimator of regression coefficient which is the OLSE in a no-measurement error linear regression model, so one idea would be to use the corrected estimators of regression coefficients in the expression of R^2. It is well known that the OLSE can be corrected for the presence of measurement error in the data by using some additional prior information available from outside the sample. In the case of multiple measurement error model, the additional information in the form of known covariance matrix of measurement errors associated with explanatory variables and known reliability matrix of explanatory variables are the two popular forms which provide the consistent estimate of regression coefficient vector, see, e.g. Cheng and Van Ness (1991), Schneeweiss (1976), Gleser (1992, 1993), Shalabh (2003) etc. Our modest objective in this paper is to use both types of available information and obtain an appropriate coefficient of determination which can be used to judge the goodness of fit of a measurement error model.

The plan of the paper is as follows. The multivariate ultrastructural model and various statistical assumptions are described in Section 2. In Section 3, we study the inconsistency of coefficient of determination under ultrastructural measurement error model. We propose two new coefficients of determination which are consistent for appropriate parameter in Section 4. Asymptotic normality of the proposed coefficients of determination is established under ultrastructural measurement error model in Section 5.

2 The model

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

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

where α is intercept term, $\mathbf{1}_n$ is a $n \times 1$ vector of elements unity ($1'$s), and β is $p \times 1$ vector of regression coefficients. We assume that η and Ξ can not be observed accurately. They are observed with additive measurement errors as

$$y = \eta + \epsilon \text{ and } X = \Xi + \Delta, \tag{2.2}$$

respectively. Here, $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)'$ and $\Delta = (\delta_1', \delta_2', \ldots, \delta_n')'$ are measurement errors associated with η and Ξ, respectively. Further, we assume that

$$\Xi = M + \Phi, \tag{2.3}$$

where $E(\Xi) = M = (\mu_1', \mu_2', \ldots, \mu_n')'$ is a $n \times p$ matrix of unknown constants and $\Phi = (\phi_1', \phi_2', \ldots, \phi_n')'$ is the random matrix. The $p \times 1$ random vectors $\phi_1, \phi_2, \ldots, \phi_n$ of matrix Φ are assumed to be distributed independently and identically 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 distributed independently and identically with mean vector 0, covariance matrix Σ_δ and finite moment up to order four. Further, the random variables $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ are assumed to be distributed independently and identically with mean zero and variances σ_ϵ^2 and finite fourth moment. The equations (2.1)-(2.3) describe the set up of an ultrastructural model (see Dolby (1976)) which is a synthesis of structural and functional forms of measurement error model. We further assume that $\lim_{n\to\infty} n^{-1}M'PM =: \Sigma_\mu$ which is a symmetric and positive definite matrix; here $P = I_n - n^{-1}\mathbf{1}_n\mathbf{1}_n'$. This assumption is needed for the validity of asymptotic results and avoids the possibility of any trend in the data, see Schneeweiss (1991).

3 Coefficient of determination in classical regression model

First we discuss the role of R^2 in the classical regression model under the usual assumptions. Let us consider the classical multiple linear regression model,

$y^* = \alpha\mathbf{1} + X^*\beta + u$, where explanatory variables are non-stochastic and measurement errors are absent. Here y^* is the $n \times 1$ vector of values on study variable, X^* is the $n \times p$ matrix of values on p non-stochastic explanatory variables, α is intercept term, β is $p \times 1$ vector of regression slopes, and u is $n \times 1$ vector of disturbances. Under this classical multiple linear regression model the coefficient of determination is defined as $R^{*2} = \frac{b^{*\prime}X^{*\prime}PX^*b^*}{y^{*\prime}Py^*} = \frac{y^{*\prime}PX^*(X^{*\prime}PX^*)^{-1}X^{*\prime}Py^*}{y^{*\prime}Py^*}$, where $b^* = (X^{*\prime}PX^*)^{-1}X^{*\prime}Py^*$ is the ordinary least squares estimator of vector regression coefficients β. It can be proved using the variance of disturbance terms as σ^2 and usual assumptions of classical linear regression analysis that $\text{plim}_{n\to\infty}(R^{*2} - \theta^*) = 0$, where $\theta^* = \frac{\beta'(n^{-1}X^{*\prime}PX^*)\beta}{\beta'(n^{-1}X^{*\prime}PX^*)\beta+\sigma^2}$ is the population counterpart of R^2 and denotes the population multiple correlation coefficient between the study and explanatory variables. Thus it is established that R^{*2} is a consistent estimator of θ^* in the absence of measurement errors.

Note that the equations (2.1) - (2.2) can jointly be written as $y = \alpha\mathbf{1}_n + \Xi\beta + \epsilon$. Under this model, the coefficient of determination can be defined on the similar lines as in the case of classical regression model without measurement errors as $R^{\dagger 2} = \frac{y'P\Xi(\Xi'P\Xi)^{-1}\Xi'Py}{y'Py}$. Let $\Sigma^\dagger := \text{plim}_{n\to\infty}n^{-1}\Xi'P\Xi$, then $\text{plim}_{n\to\infty}(R^2 - \theta) = 0$, where $\theta^\dagger = \frac{\beta'\Sigma^\dagger\beta}{\beta'\Sigma^\dagger\beta+\sigma_\epsilon^2}$. So this establishes a sort of similarity of $R^{\dagger 2}$ between the linear regression models with and without measurement errors.

4 Coefficient of determination in measurement error model

Unfortunately in measurement error models, Ξ is not observable and can only be observed as X with measurement errors given by Δ. So we replace the unobservable Ξ by observable X in the expression of R^2 and obtain the expression of coefficient of determination as $R^2 = \frac{y'PX(X'PX)^{-1}X'Py}{y'Py}$. In case X has no measurement errors, the coefficient of determination defined R^2 is consistent for estimating the parameter $\theta = \frac{\beta'\Sigma\beta}{\beta'\Sigma\beta+\sigma_\epsilon^2}$, where $\Sigma := \text{plim}_{n\to\infty}n^{-1}X'PX$. It is well known that the ordinary least squares estimate of regression coefficients becomes inconsistent in the presence of measurement errors in the data. In order to estimate the regression slopes consistently, some additional information about unknown parameters is required. For this purpose various types of information are discussed in the literature, e.g., the covariance matrix of measurement errors associated with explanatory variables, reliability matrix associated with explanatory variables, the availability of observations on some instrumental variables etc. Therefore, there is a doubt on the consistency of R^2 for the parameter θ in the presence of measurement errors in the model. We below present some lemma which are useful in proving the inconsistency of R^2.

Lemma 1 *Under the model (2.1)-(2.3) and the assumptions made in Section 2,*

(i) $plim_{n\to\infty}n^{-1}X'PX = \Sigma_\mu + \Sigma_\phi + \Sigma_\delta = \Sigma,$

(ii) $plim_{n\to\infty}n^{-1}X'Py = (\Sigma - \Sigma_\delta)\beta,$

(iii) $plim_{n\to\infty}n^{-1}y'Py = \beta'(\Sigma - \Sigma_\delta)\beta + \sigma_\epsilon^2,$

(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 *Using the results of Lemma 1, we have*

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

The theorem can be easily proved using Lemma 1. Clearly, under measurement error models, $\text{plim}_{n\to\infty} R^2 \neq \theta$, in general which proves that R^2 is an inconsistent estimator of θ.

In order to obtain new coefficients of determination, 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 vector 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.

4.1 Σ_δ is known

In order to obtain an improved and modified form of the coefficient of determination which is consistent for estimating the parametric function $\theta = \frac{\beta'\Sigma\beta}{\beta'\Sigma\beta+\sigma_\epsilon^2}$, we look for the consistent estimates of $\beta'\Sigma\beta$ and σ_ϵ^2. When Σ_δ is known, the consistent estimator of β in ultrastructural model (2.1)-(2.3) is $b_\delta = (S - \Sigma_\delta)^{-1}Sb$, where $S = n^{-1}X'PX$, see Cheng and Van Ness (1999) and Fuller (1987). Using Lemma 1(i), b_δ, and the fact $\text{plim}_{n\to\infty}\{n^{-1}y'Py - b_\delta'(S - \Sigma_\delta)b_\delta\} = \sigma_\epsilon^2$, we propose a new coefficient of determination under the knowledge of Σ_δ, as $R_\delta^2 = \frac{b_\delta' S b_\delta}{b_\delta' S b_\delta + \{n^{-1}y'Py - b_\delta'(S-\Sigma_\delta)b_\delta\}} = \frac{b_\delta' S b_\delta}{n^{-1}y'Py + b_\delta'\Sigma_\delta b_\delta}$, provided $b_\delta' S b_\delta \geq n^{-1}y'Py + b_\delta'\Sigma_\delta b_\delta$. 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. It can be easily seen that $\text{plim}_{n\to\infty} R_\delta^2 = \theta$. Thus, it is clear that the proposed coefficient of determination R_δ^2 is a better choice as a measure of goodness of fit of a linear regression model in the presence of measurement errors.

4.2 Reliability matrix is known

When reliability matrix $K_x = \Sigma_x^{-1}(\Sigma_x - \Sigma_\delta)$ is known, the consistent estimator of β in ultrastructural model (2.1)-(2.3) is $b_k = K_x^{-1}b$, see Cheng and Van Ness (1999) and Fuller (1987). This estimator has its own advantages, see Gleser (1992, 1993) for more details. Using Lemma 1(i), b_k, and the fact $\text{plim}_{n\to\infty}\{n^{-1}y'Py - b_k' S K_x b_k\} = \sigma_\epsilon^2$, we propose a new coefficient of determination under the knowledge of K_x , as $R_k^2 = \frac{b_k' S b_k}{b_k' S b_k + \{n^{-1}y'Py - b_k' S K_x b_k\}} = \frac{b_k' S b_k}{n^{-1}y'Py + b_k' S(I_p - K_x)b_k}$, provided $b_k' S b_k \geq n^{-1}y'Py + b_k' S(I_p - K_x)b_k$. 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. Using the results of Lemma 1, and the consistency of b_k for estimating β, it can be easily proved that $\text{plim}_{n\to\infty} R_k^2 = \theta$. Thus, it is clear that the proposed coefficient of determination R_k^2 is a better choice as a measure of goodness of fit of a linear regression model in the presence of measurement errors.

5 Asymptotic properties

For the sake of simplicity, we assume, without loss of generality, that all the data are mean corrected and $\Sigma_\delta = \sigma_\delta^2 I_p$ and $\Sigma_\phi = \sigma_\phi^2 I_n$. 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, for $i=1,2,\ldots,n, j=1,2,\ldots,p$, ϵ_i, δ_{ij} and ϕ_{ij} are also assumed to be statistically independent. We further assume that the nth row of matrix M converges to σ_μ'. Consequently, we have $\lim_{n\to\infty} n^{-1}M'PM = \lim_{n\to\infty} n^{-1}M'M = \sigma_\mu\sigma_\mu' = \Sigma_\mu$.

Lemma 2 *Define*

$$\begin{aligned}
H &= \sqrt{n}(S-\Sigma_x), && (5.1)\\
h &= \sqrt{n}\{n^{-1}X'(\epsilon-\Delta\beta)+\sigma_\delta^2\beta\}, && (5.2)\\
g &= \sqrt{n}\{n^{-1}(\epsilon-\Delta\beta)'(\epsilon-\Delta\beta)-\sigma_\epsilon^2-\sigma_\delta^2\beta'\beta\}, && (5.3)
\end{aligned}$$

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

$$\begin{aligned}
\Omega_H(dd') &= (\sigma_\delta^2+\sigma_\phi^2)[\Sigma\{dd'+(d'd)I_p\}+dd'\sigma_\mu\sigma_\mu'+(d'\sigma_\mu\sigma_\mu'd)I_p]\\
&\quad +(\gamma_{1\phi}\sigma_\phi^3+\gamma_{1\delta}\sigma_\delta^3)[f(\sigma_\mu\mathbf{1}_p',dd')+\{f(\sigma_\mu\mathbf{1}_p',dd')\}'\\
&\quad +2f(I_p,\mathbf{1}_p\sigma_\mu dd')]+(\gamma_{2\phi}\sigma_\phi^4+\gamma_{2\delta}\sigma_\delta^4)f(I_p,dd'),\\
\Omega_{hH}(d) &= -\sigma_\delta^2[\Sigma(d\beta'+(d'\beta)I_p)+\gamma_{1\delta}\sigma_\delta\{f(\sigma_\mu\mathbf{1}_p',\beta d')\\
&\quad +f(I_p,\beta d'\sigma_\mu\mathbf{1}_p)+(f(\sigma_\mu\mathbf{1}_p',d\beta'))'\}+\gamma_{2\delta}\sigma_\delta^2 f(I_p,\beta d')],\\
\Omega_h &= (\sigma_\epsilon^2+\sigma_\delta^2(\beta'\beta))\Sigma+\sigma_\delta^4\beta\beta'+\gamma_{1\delta}\sigma_\delta^3\{f(\sigma_\mu\mathbf{1}_p',\beta\beta')\\
&\quad +\big(f(\sigma_\mu\mathbf{1}_p',\beta\beta')\big)'\}+\gamma_{2\delta}\sigma_\delta^4 f(I_p,\beta\beta'),\\
\Omega_g &= 2(\sigma_\epsilon^2+\sigma_\delta^2\beta'\beta)^2+\gamma_{2\epsilon}\sigma_\epsilon^4+\gamma_{2\delta}\sigma_\delta^4 f(I_p,\beta\beta')\beta,\\
\Omega_{gH}(d) &= \big[\gamma_{1\delta}\sigma_\delta^3\{f(\sigma_\mu\mathbf{1}_p,d\beta')+f(I_p,\beta d'\sigma_\mu\mathbf{1}_p)\}+\gamma_{2\delta}\sigma_\delta^4 f(I_p,d\beta')\\
&\quad +\sigma_\delta^2\beta d'+\sigma_\delta^4\beta' d\big]\beta\\
\Omega_{gh} &= -2\sigma_\epsilon^2\sigma_\delta^2\beta-2\sigma_\delta^4\beta\beta'\beta+\gamma_{1\epsilon}\sigma_\epsilon^3\sigma_\mu-\gamma_{1\delta}\sigma_\delta^3 f(\sigma_\mu\mathbf{1}_p',\beta\beta')-\gamma_{2\delta}\sigma_\delta^4 f(I_p,\beta\beta').
\end{aligned}$$

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* $Z_1, Z_2\in\mathbb{R}^{p\times p}$, *where* $*$ *denotes the Hadamard product operator of matrices and* $\xrightarrow{d}$ *denotes the convergence in distribution.*

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 on replacing d by β in the covariance matrix given in Lemma 2.

Theorem 3 *As $n \longrightarrow \infty$, $\sqrt{n}\left(R_k^2 - \frac{\beta'\Sigma_x\beta}{\beta'\Sigma_x\beta+\sigma_\epsilon^2}\right) \xrightarrow{d} N\left(0, \omega_k'\Omega_k\omega_k\right)$,*

where $\omega_k = \begin{pmatrix} \sigma_\epsilon^2\beta \\ (\Sigma K_x)^{-1}\{\sigma_\epsilon^2 I_p - \bar{K}_x\Sigma\beta'\beta\}\Sigma\beta \\ (\Sigma K_x)^{-1}\Sigma\beta(\sigma_\epsilon^2 - \beta'\Sigma\beta) \\ \beta'\Sigma\beta \end{pmatrix}$ *and*

$$\Omega_k = \begin{pmatrix} \Omega_H(\beta\beta') & \Omega_H(\beta\beta'\bar{K}_x) & \Omega_{hH}(\beta) & \Omega_{gH}(\beta) \\ \Omega_H'(\beta\beta'\bar{K}_x) & \Omega_H(\bar{K}_x\beta\beta'\bar{K}_x) & \Omega_{hH}(\bar{K}_x\beta) & \Omega_{gH}(\bar{K}_x\beta) \\ \Omega_{hH}'(\beta) & \Omega_{hH}'(\bar{K}_x\beta) & \Omega_h & \Omega_{gh} \\ \Omega_{gH}'(\beta) & \Omega_{gH}'(\bar{K}_x\beta) & \Omega_{gh}' & \Omega_g \end{pmatrix},$$

$\bar{K}_x = (I_p - K_x)$. The elements of the asymptotic covariance matrix Ω_k are obtained using Lemma 2.

References

Cheng, C. L. and Van Ness, J.W. (1991): "On the unreplicated ultrastructural model" *Biometrika*, 78, 442-445.

Cheng C.-L. and Van Ness J. W. (1999): *Statistical Regression with Measurement Errors*, London, Arnold.

Cramer, J. S. (1987): "Mean and variance of R^2 in small and moderate samples" *Journal of Econometrics*, 35, 253-266.

Dolby G. R. (1976): "The ultrastructural relation: a synthesis of the functional and structural relations" *Biometrika*, 63, 39-50.

Fuller, W. A. (1987): *Measurement Error Models*, John Willey.

Gleser, L. J. (1984): "A note on G.R. Dolby's ultrastructural model" *Biometrika*, 72, 117-124.

Gleser, L. J. (1992): "The Importance of assessing measurement reliability in multivariate regression" *Journal of American Statistical Association*, 87(419), 696-707.

Gleser, L. J. (1993): "Estimators of slopes in linear errors-in-variables regression models when the predictors have known reliability matrix" *Statistics and Probability Letters*, 17, 113-121 .

Nagelkerke, N. J. D. (1991): "A note on a general definition of the coefficient of determination" *Biometrika*, 78(3), 691-692.

Ohtani K. (1994): "The density functions of R^2 and $\bar{R}^2$, and their risk performance under asymmetric loss in misspecified linear regression models" *Economic Modelling*, 11(4), 463-471.

Schneeweiss, H. (1976): "Consistent Estimation of a regression with Errors in the variables" *Metrika*, 23, 101-115.

Schneeweiss, H. (1991): "Note on a linear model with errors in the variables and with trend" *Statistical Papers*, 32, 261-264.

Shalabh (2003). "Consistent estimation of coefficients in measurement error models under non-normality" *Journal of Multivariate Analysis*, 86(2), 227-241.

Srivastava, A. K., Shobhit (2002): "A family of estimators for the coefficient of determination in linear regression models" *Journal of Applied Statistical Science*, 11(2), 133-144.

Srivastava, A.K., Srivastava V.K. and Ullah A. (1995): "The coefficient of determination and its adjusted version in linear regression models" *Econometric Reviews*, 14, 229-240.