Теорія ймовірностей
та математична статистика
Вип. 1(100)/2019, с. 177–197

Teoriya Ĭmovirnosteĭ
ta Matematychna Statystyka
No. 1(100)/2019, pp. 177–197

UDC 519.21

GOODNESS OF FIT FOR GENERALIZED SHRINKAGE ESTIMATION

C.-L. CHENG, SHALABH, A. CHATURVEDI

Abstract. The present paper develops a goodness of fit statistic for the linear regression models fitted by the shrinkage type estimators. A family of double k-class estimators is considered as a shrinkage estimator which encompasses several estimators as its particular case. The covariance matrix of error term is assumed to be a non-identity matrix under two situations- known and unknown. The goodness of fit statistics based on the idea of coefficient of determination in multiple linear regression model is proposed for the family of double k-class estimators. Its first and second order moments up to the first order of approximation are derived and finite sample properties are studied using the Monte-Carlo simulation.

Key words and phrases. Linear regression, non-spherical disturbances, coefficient of determination (R^2), shrinkage estimation, generalized least squares estimator, feasible double k-class estimators, feasible generalized least squares estimator, double k-class estimators, goodness of fit.

2010 *Mathematics Subject Classification.* 62J07; 62J05.

1. Introduction

Various estimation procedures and estimators are available in the literature to estimate the regression parameters in a multiple linear regression model. The ordinary least squares estimation procedure provides the best linear unbiased estimator of the regression coefficient under the Gauss Markov setup. If the linearity and unbiasedness properties of an estimator can be compromised, then the shrinkage estimation provides an estimator of regression coefficients which is more efficient than the ordinary least squares estimator under some mild conditions. For example, the family of Stein rule estimators provides such estimators which have smaller variance than the ordinary least squares estimator under a mild condition that the number of explanatory variables are more than two. Motivated by the concept from shrinkage estimation, [3] proposed the family of double k-class estimators for estimating the regression coefficients. Such an estimator has a very general form and encompasses several prominent estimators proposed in the literature as its particular case. Later [3] (see also, [4]) derived the exact and large non-centrality parameters approximations of bias, mean squared error matrix and risk of double k-class estimators. The properties of double k-class estimators have been investigated by various authors in the literature, see [2, 6] and [17] for details on the developments for the double k-class estimators.

The double k-class estimators are characterized by two characterizing scalars. Substituting different values of these scalars, one can obtain various other estimators. In a more general set up, if the error distribution is non spherical and the covariance matrix of non spherical errors is unknown, various forms of feasible versions of double k-class estimators are proposed in the literature see [2] and [6] for a comprehensive presentation about these estimators.

All such estimators are obtained through different procedures and have different properties. Usually an experimenter is more interested in getting a well fitted model rather than knowing the performance of these estimators when they are actually used in fitting a linear regression model. Suppose a linear regression model is fitted using two different choices of characterizing scalars. The choices of such scalars are optimal under their own but different criterion. How it affects the model fitting or how to know which choice of scalars gives a better fitted model remains unclear in such situations. Even in the case the experimenters use different sets of variables, the question about which model is better remains unanswered. Thus an important question arises: how to conduct the goodness-of-fit in the models which are fitted using the double k-class estimators?

The goodness-of-fit in the usual multiple linear regression model is usually judged by the coefficient of determination (COD) which is the multiple correlation coefficient between the study variable and all the explanatory variables in the model. It is usually denoted by R^2. In the literature, several researchers have attempted to obtain the forms of COD for linear models under nonstandard conditions, see [7] (see also [8]) for detailed review on the development and research done on COD in various type of linear models.

The R^2 statistic is based on the ordinary least squares estimates of regression coefficients which is the best linear unbiased estimator. It is obtained by the analysis of variance in linear regression model which is again based on the ordinary least squares estimates. The R^2 is a statistic and is a biased but consistent estimator of population multiple correlation coefficient in the set up of multiple linear regression model, see Chapter 4 in [19]. The double k-class estimators of regression coefficient are not the best linear unbiased estimators of regression coefficient, rather they are nonlinear and biased estimators of regression coefficient. So conducting the analysis of variance based on the double k-class estimators in the same way as with the ordinary least squares estimator is not possible. Hence the COD can not be obtained in the same way as with ordinary least squares estimates in such a case and so the bigger question arises: how to judge the goodness-of-fit in such a case. Moreover, any goodness-of-fit statistic is itself a statistic, i.e., a function random variable which is being used to estimate any unknown parameter. For example, R^2 in case of multiple linear regression model is a statistic whose statistical properties are determined by its moments. Such moments help in finding the confidence interval, testing of hypothesis etc.

We have attempted in this direction. We have considered the estimation of regression coefficients using the double k-class estimators. A goodness-of-fit statistic is proposed to judge the quality of such fitted model. The first and second order moments of such statistic up to the first order of approximation are derived. The finite sample properties and performance of double k-class estimators are studied through the Monte-Carlo simulation.

The plan of the paper is as follows. Section 2 describes the linear model and the double k-class estimators along with its feasible version. The goodness-of-fit statistics are developed in Section 3. The moments of the goodness of fit statistics are derived in Section 4. The results from the Monte-Carlo simulation are reported in Section 5 and some conclusions are presented in Section 6.

An extended version of the paper with more details on earlier work and simulations is available at `http://home.iitk.ac.in/~shalab/r2dkcsc.pdf`.

2. The Model and Estimators

Consider the following linear regression model between an $n \times 1$ vector y of n observations of the study variable and an $n \times p$ matrix X of n observations of p explanatory variables:

$$y = \alpha \mathbf{1}_n + X\beta + \epsilon, \tag{1}$$

where α is the intercept term, $\mathbf{1}_n$ is an $n \times 1$ vector with unit elements ($1'$s), β is a $p \times 1$ vector of coefficients associated with them and ϵ is an $n \times 1$ vector of non-spherical disturbances. The disturbance vector ϵ is assumed to follow a multivariate normal distribution with mean vector 0 and known variance-covariance matrix $\sigma^2\Omega^{-1}$.

Define $A = \Omega - \frac{1}{\mathbf{1}_n'\Omega\mathbf{1}_n}\Omega\mathbf{1}_n\mathbf{1}_n'\Omega$ and pre-multiplying (1) by A gives the model

$$Ay = AX\beta + A\epsilon\,. \tag{2}$$

The generalized least squares estimator (GLSE) to estimate β in (2) gives the best linear unbiased estimator of β as

$$\hat{\beta}_g = (X'AX)^{-1}X'Ay. \tag{3}$$

In the spirit of double k-class estimator proposed by [3] (see also [4]), the family of generalized double k-class estimators (DKKE) characterized by two nonstochastic scalars $k_1 > 0$ and $0 < k_2 < 1$ is given as

$$\begin{aligned}\hat{\beta}_{kk} &= \left[1 - \left(\frac{k_1}{n-p+2}\right)\frac{(y-X\hat{\beta}_g)'A(y-X\hat{\beta}_g)}{y'Ay - k_2(y-X\hat{\beta}_g)'A(y-X\hat{\beta}_g)}\right]\hat{\beta}_g = \\ &= \left[1 - \left(\frac{k_1}{n-p+2}\right)\frac{v}{\hat{\beta}_g'X'AX\hat{\beta}_g + (1-k_2v)}\right]\hat{\beta}_g,\end{aligned} \tag{4}$$

where $v = (y - X\hat{\beta}_g)'A(y - X\hat{\beta}_g)$. This is a very general class of estimator which gives rise to various estimators considered in the literature as the special case of DKKE. For example, $k_1 = 0$ gives GLSE; $k_1 = p - 2$ and $k_2 = 1$ gives generalized Stein-rule estimator (GSRE); $k_1 = \frac{1}{n-p}$ and $k_2 = 1 - k_1$ gives generalized minimum mean squared error estimator (GMMSE); $k_1 = \frac{n-p+2}{n-p}$ and $k_2 = 1 - \frac{k_1}{n-p+2}$ gives adjusted generalized minimum mean squared error estimator (AGMMSE); $k_1 = \frac{(n-p+2)p}{n-p}$ and $k_2 = 1 - \frac{k_1}{n-p+2}$ gives generalized double k-class estimators (GKKCE) etc.

In case, the variance-covariance matrix $\sigma^2\Omega^{-1}$ is unknown, we assume that the elements of Ω are functions of a $q \times 1$ parameter vector θ that belongs to an open subset of q dimensional Euclidean space. We also assume that a consistent estimator $\hat{\theta}$ of θ is available so that Ω is consistently estimated by $\hat{\Omega} = \Omega(\hat{\theta})$. So $\Omega \equiv \Omega(\theta)$ and $\hat{\Omega} \equiv \Omega(\hat{\theta})$. In such situation, β in (2) is estimated by the feasible generalized least squares estimator (FGLSE) of β given by

$$\hat{\beta}_{fg} = (X'\hat{A}X)^{-1}X'\hat{A}y. \tag{5}$$

In the spirit of [3] (see also [4]) and [2], the family of feasible generalized double k-class estimators (FDKKE) is defined by

$$\begin{aligned}\hat{\beta}_{fkk} &= \left[1 - \left(\frac{k_1}{n-p+2}\right)\frac{(y-X\hat{\beta}_{fg})'\hat{A}(y-X\hat{\beta}_{fg})}{y'\hat{A}y - k_2(y-X\hat{\beta}_{fg})'\hat{A}(y-X\hat{\beta}_{fg})}\right]\hat{\beta}_{fg} = \\ &= \left[1 - \left(\frac{k_1}{n-p+2}\right)\frac{\hat{v}}{y'\hat{A}y - k_2\hat{v}}\right]\hat{\beta}_{fg} = \\ &= \left[1 - \left(\frac{k_1}{n-p+2}\right)\frac{\hat{v}}{\hat{\beta}_{fg}'X'\hat{A}X\hat{\beta}_{fg} + (1-k_2)\hat{v}}\right]\hat{\beta}_{fg},\end{aligned} \tag{6}$$

where $\hat{v} = (y - X\hat{\beta}_{fg})'\hat{A}(y - X\hat{\beta}_{fg})$ and $\hat{A}$ is obtained by replacing Ω by $\hat{\Omega}$ in A. Again, FDKKE gives rises to various estimators considered in the literature as its special cases which are obtained by substituting the values of k_1 and k_2. For example, $k_1 = 0$ gives feasible generalized least squares estimator (FGLSE); $k_1 = p - 2$ and $k_2 = 1$ gives feasible generalized Stein-rule estimator (FGSRE); $k_1 = \frac{1}{n-p}$ and $k_2 = 1 - k_1$ gives feasible generalized minimum mean squared error estimator (FGMMSE), proposed by [16], see also [10, 11]; $k_1 = \frac{n-p+2}{n-p}$ and $k_2 = 1 - \frac{k_1}{n-p+2}$ gives adjusted feasible generalized minimum mean squared error estimator (AFGMMSE), proposed by [11]; $k_1 = \frac{(n-p+2)p}{n-p}$

and $k_2 = 1 - \frac{k_1}{n-p+2}$ gives feasible generalized double k-class estimator (FGKKCE), proposed by [15].

It is clear from the description of DKKE and FDKKE that these families encompass several popular estimators as their particular cases. Use of these estimators gives rise to different fitted models. Researchers have tried to obtain the optimal values of characterizing scalars under certain criterion, e. g., mean squared error of estimators, asymptotic variances etc. Even prominent results are based on certain approximations like asymptotic approximations, small error approximations etc. If we try looking at the modelling issue from the practitioner's point of view, then the practitioner is ultimately more interested in knowing whether the fitted model is good or not and if it can be used for other purposes like forecasting? Moreover, the practitioner is working with finite data. The use of an asymptotically efficient estimator may not necessarily ensure that the fitted model obtained with finite data is also superior. Also, if two different models are fitted using the same set of data but different transformations on either study variable, explanatory variables or both, then it can not be determined which model provides better fit. So we need to have a goodness-of-fit statistic which can help the practitioner to know whether the obtained fitted model is good enough to work further or not. We attempt to develop goodness-of-fit statistics in the next section.

3. goodness-of-fit Statistics (GoFs) for Estimators

Our objective is to define a goodness-of-fit statistic for the family of DKKE and FDKKE. Since the COD in the classical linear regression model is a popular statistic to judge the goodness of fit, we borrow the idea from there and use it do develop a suitable statistic. First, we develop the goodness-of-fit statistic based on the COD based on GLSE of regression coefficient.

3.1. **Development and Consistency of GoF for GLSE.** Consider the form of R^2 in the regression model (2) under the assumptions $V(\epsilon) = \sigma^2\Omega^{-1} = \Sigma^*$, which is a positive definite matrix. One can find a nonsingular matrix K such that $KK^{-1} = \Sigma^*$. Pre-multiplying (2) by K^{-1} gives

$$K^{-1}Ay = K^{-1}AX\beta + K^{-1}A\epsilon\,, \tag{7}$$

where $y^* = K^{-1}Ay$, $X^* = K^{-1}AX$ and $\epsilon^* = K^{-1}A\epsilon$. Note that now $E(\epsilon^*) = 0$ and $V(\epsilon^*) = \sigma^2 I$ so that elements in ϵ^* are identically and independently normally distributed. The goodness of fit statistic based on the idea from COD under (7) is proposed as

$$R_g^2 = \frac{y^{*\prime}X^*(X^{*\prime}X^*)^{-1}X^{*\prime}y^*}{y^{*\prime}y^*} = \frac{y'AX(X'AX)^{-1}X'Ay}{y'Ay} = \frac{\hat{\beta}_g'X'AX\hat{\beta}_g}{y'Ay}\ ,\ \ 0 \le R_g^2 \le 1. \tag{8}$$

Assume that

$$\operatorname*{plim}_{n\to\infty} \frac{X'AX}{n} = \Sigma_{XX}\ \text{(a positive definite matrix), and}\ \operatorname*{plim}_{n\to\infty} \frac{X'A\epsilon}{n} = 0 \tag{9}$$

where the notation plim denotes the limit in probability.

The COD is essentially the square of sample based multiple correlation coefficient between the study variable and all the explanatory variables in the linear regression model. The conventional population based multiple correlation coefficient is given by

$$\vartheta = \frac{\beta'\Sigma_{XX}\beta}{\beta'\Sigma_{XX}\beta + \sigma^2}\ \ ;\ \ 0 \le \vartheta \le 1, \tag{10}$$

assuming A to be known, see Chapter 4 in [19]. If the model is best fitted, then $\sigma^2 = 0$ which implies $\vartheta = 1$. If the model is worst fitted in the sense that no explanatory variable contributes in the modelling, then $\beta = 0$ which in turn implies $\vartheta = 0$. Any other values of $0 \leq \vartheta \leq 1$ indicate the degree of GoF of the model explained by ϑ. So ϑ acts as a measure of goodness-of-fit in the sense of contribution of explanatory variables to the explanation of the variability among the values of study variable obtained by the model.

Using (9), it can be shown that $\text{plim}_{n\to\infty} \hat{\beta}_g = \beta$, where $\text{plim}_{n\to\infty}$ denotes the convergence in probability and consequently, such defined R_g^2 in (8) turns out to be a consistent estimator of ϑ in the sense that $\text{plim}_{n\to\infty} R_g^2 = \vartheta$, where R_g^2 is the sample counterpart of the population multiple correlation coefficient ϑ.

Note that $0 \leq R_g^2 \leq 1$. When all the explanatory variables in the model are not contributing towards the explanation of variation of values of the study variable, then ideally $\hat{\beta}_g$ will be zero or close to zero. In turn, $R_g^2 = 0$ indicating that the model fit is poor. Similarly, $R_g^2 = 1$ will indicate the perfect fit. Any other value of R_g^2 between 0 and 1 will give an idea about the degree of goodness-of-fit in the model, similar to COD in a usual multiple linear regression model. So R_g^2 in (8) defines a statistic like the COD for the linear regression model having a non-identity covariance matrix assuming A (or equivalently Ω) is known. This statistic can be used to judge the goodness-of-fit of a model when GLSE is used to fit the model.

3.2. **Development and Consistency of GoFs for DKKE.** Next, we propose using $\hat{\beta}_{kk}$ in place of $\hat{\beta}_g$ in (8) to obtain a statistic given as

$$R_{kk}^2 = \frac{\hat{\beta}'_{kk} X' A X \hat{\beta}_{kk}}{y' A y} \quad , \quad 0 \leq R_{kk}^2 \leq 1. \tag{11}$$

Note that $0 \leq R_{kk}^2 \leq 1$, so this also has an interpretation like the COD in a multiple linear regression model. For example, when $\hat{\beta}_{kk}$ is zero or close to zero, it indicates that all the elements in its vector are zero or close to zero meaning thereby that all the explanatory variables in the model are not capable of explaining the variation in the values of a study variable. Hence $R_{kk}^2 = 0$ will indicate the poorest fit. Similarly, $R_{kk}^2 = 1$ will indicate the best fit and any other value of R_{kk}^2 between 0 and 1 will give an idea about the degree of goodness-of-fit in the model resulting by the use of any estimator from the family of DKKE. Thus, this statistic can be used to judge the goodness-of-fit of a model based on DKKE $\hat{\beta}_{kk}$.

Using (9), we can show that $\text{plim}_{n\to\infty} \beta_{kk} = \beta$ and consequently $\text{plim}_{n\to\infty} R_{kk}^2 = \vartheta$, $0 \leq \vartheta \leq 1$. Thus, it is established that R_{kk}^2 is a consistent estimator of the population multiple correlation coefficient ϑ.

It may be noted that R_g^2 and R_{kk}^2 are biased estimators of ϑ.

3.3. **Development and Consistency of GoFs for FGLSE and FDKKE.** Now we consider the case when A (or equivalently Ω) is unknown and is consistently estimated as $\hat{A}$ (or equivalently $\hat{\Omega}$) such that $\text{plim}_{n\to\infty} \hat{A} = A$ (or equivalently $\text{plim}_{n\to\infty} \hat{\Omega} = \Omega$). In such a case, the GLSE and DKKE are estimated by (5) and (6) respectively. In such case, the goodness-of-fit statistics are proposed as

$$R_{fg}^2 = \frac{\hat{\beta}'_{fg} X' \hat{A} X \hat{\beta}_{fg}}{y' \hat{A} y} \quad , \quad 0 \leq R_{fg}^2 \leq 1 \tag{12}$$

and

$$R_{fkk}^2 = \frac{\hat{\beta}'_{kk} X' \hat{A} X \hat{\beta}_{kk}}{y' \hat{A} y} \quad , \quad 0 \leq R_{fkk}^2 \leq 1, \tag{13}$$

respectively.

Using (9), we can show that $\text{plim}_{n\to\infty} R_{fg}^2 = \vartheta$ and $\text{plim}_{n\to\infty} R_{fkk}^2 = \vartheta$, $0 \leq \vartheta \leq 1$. Thus, both R_{fg}^2 and R_{fkk}^2 are again consistent but biased estimator of the population multiple correlation coefficient ϑ. Both R^2_{fg} and R^2_{fkk} lie between 0 and 1. The interpretations of R^2_{fg} and R^2_{fkk} can be obtained just like the interpretations of R^2_g and R^2_{kk}, respectively. For example, $R^2_{fg} = 0$ and $R^2_{fkk} = 0$ indicate the poorest fit, and $R^2_{fg} = 1$ and $R^2_{fkk} = 1$ indicate the best fit of the model using the estimators $\hat{\beta}_{fg}$ and $\hat{\beta}_{kk}$, respectively.

3.4. **Earlier Studies about COD.** The properties of COD have been studied in the literature when the covariance matrix of disturbance is the identity matrix, i.e., disturbances are identically and independently distributed. Under such an assumption, Crämer [9] studied the properties of COD and adjusted COD under normally distributed disturbances. Ohtani and Giles [12] considered the criterion of risk under an absolute error loss function and studied the relative performance of COD and its adjusted version. Ohtani and Hasegawa [13] analyzed the properties of COD and its adjusted version under the multivariate t-distribution of disturbances in the misspecified model. The exact expressions of the properties of COD and/or its adjusted version in all such studies turn out to be complicated and so it is very difficult to draw any clear conclusions. Some inferences are drawn based on numerical computations for some choice of parameters and so they have very limited utility in applications. So some approximations of the statistical properties may be more useful to shed some light on the utility of the COD. Such an idea is used in [5] where they have utilized the small disturbance asymptotic theory to obtain the approximate moments of COD. Such results are valid only when the fitted model is close to perfect fit. This constraint diminishes the utility of the statistical inferences in real applications. Smith [14] studied the closeness of exact moments derived by using the small disturbance and other asymptotic approximations. Alternatively, the asymptotic theory can also be utilized to derive the asymptotic distribution and to draw statistical conclusions. It has an advantage that it does not require the assumption of any specific distribution of errors like normal. When COD is derived in the usual linear regression model, it is not based on the assumption of normality of disturbances. So use of asymptotic theory to study the properties of COD has an advantage. Srivastava et al. [1] derived the bias and mean squared errors of COD and its adjusted version using the large sample asymptotic theory. We also propose to use the large sample asymptotic theory to derive the first and second moments of the proposed goodness-of-fit statistics. Cheng et al. [7] proposed the goodness-of-fit statistic based on the COD in measurement error models.

4. Moments of goodness-of-fit Statistics

The statistic R^2_{fkk} is of more general form than R^2_{kk} and more useful in application data. So we state and derive some asymptotic results about R^2_{fkk} in the Theorem 1.

We assume that the elements of Ω are functions of a $q \times 1$ parameter vector θ that belongs to an open subset of q-dimensional Euclidean space and a consistent estimator $\hat{\theta}$ of θ is available. Let $A = \Omega - \frac{1}{\mathbf{1}_n'\Omega\mathbf{1}_n}\Omega\mathbf{1}_n\mathbf{1}_n'\Omega$, $\Omega_i = \frac{\partial}{\partial\theta_i}\Omega$, $\Omega_{ij} = \frac{\partial^2}{\partial\theta_i\partial\theta_j}\Omega$, $B = \frac{1}{n}X'AX$, $A_j = \frac{\partial}{\partial\theta_j}A$, $B_j = \frac{1}{n}X'A_jX$, $B_{jk} = \frac{1}{n}X'A_{jk}X$, $\phi = \beta'B\beta + \sigma^2$, $P_j = \frac{1}{\sqrt{n}}\left(X'A_j - B_jB^{-1}X'A\right)$ and $\xi = \beta'B\beta + (1-k_2)\sigma^2$.

For the validity of large sample approximations for the moments, following [18], we assume that

1. (1) $\hat{\theta}$ has a stochastic expansion of the form $\sqrt{n}(\hat{\theta} - \theta) = d + o_p(n^{-1})$, where d has an asymptotic distribution $N(0, \Lambda)$, with $\Lambda = ((\lambda_{ij}))$.
2. (2) As $n \to \infty$, $\frac{1}{n}X'\Omega X$ tends to a finite positive definite matrix, the elements of $\frac{1}{n}X'\Omega_iX$, $\frac{1}{n}X'\Omega_{ij}X$, and $\frac{1}{n}X'\Omega_i\Omega^{-1}\Omega_jX$ approach to finite limits for all $i, j = 1, 2, \ldots, q$.

(3) The elements of $\frac{1}{\sqrt{n}}X'(\hat{\Omega}-\Omega))\epsilon$, $\frac{1}{n}X'(\hat{\Omega}-\Omega)X$, $\frac{1}{\sqrt{n}}X'(\hat{\Omega}_i-\Omega_i)\epsilon$, $\frac{1}{n}X'(\hat{\Omega}_i-\Omega_i)X$ converge to zero in probability.

Now, we state following two lemmas which are useful in finding the results stated in Theorem 1.

Lemma 1. *Let* $e=\sqrt{n}(\hat{\theta}-\theta)$*, where* e_j *is the* j^{th} *element of* e*,* $\eta_0=\frac{1}{\sqrt{n}}B^{-1}X'Au$*, and* $\eta_{-\frac{1}{2}}=\frac{1}{\sqrt{n}}\sum_j B^{-1}P_jue_j$*. Then we obtain the following results using the distributional properties of disturbances:* $E(e_j)=0$*,* $E(e_je_k)=\lambda_{jk}$*,* $E(\eta_0)=0$*,* $E(\eta_{-\frac{1}{2}})=0$*,* $E(\eta_0\eta_0')=\frac{\sigma^2}{n}B^{-1}X'AXB^{-1}=\sigma^2B^{-1}$*, and* $E(\eta_{-\frac{1}{2}}'B\eta_{-\frac{1}{2}})=\frac{\sigma^2}{n}\sum_{j,k}\mathrm{tr}\big(B^{-1}P_jAP_k'\big)\lambda_{jk}$*, where*

$$\begin{aligned}P_jAP_k'&=\frac{1}{n}(XA_j-B_jB^{-1}X'A)A(A_kX'-AXB^{-1}B_k)=\\&=\frac{1}{n}X'A_jAA_kX-\frac{1}{n}X'A_jAXB^{-1}B_k-\frac{1}{n}B_kB^{-1}X'AA_kX+\\&+\frac{1}{n}B_jB^{-1}X'AXB^{-1}B_k.\end{aligned}$$

Lemma 2. *Let*

$$\rho_{-\frac{1}{2}}=\frac{2}{\sqrt{n}}\beta'B\eta_0+\frac{1}{\sqrt{n}}\sum_j\beta'B_j\beta e_j-\frac{\beta'B\beta}{\phi\sqrt{n}}\beta'B\eta_0-\frac{1}{\phi\sqrt{n}}\beta'B\beta\sum_j\beta'B_j\beta e_j, \tag{14}$$

$$\begin{aligned}\rho_{-1}=\beta'B\beta\Bigg(&\frac{4}{n\phi^2}(\beta'B\eta_0)^2+\frac{1}{n\phi^2}\sum_{j,k}\beta'B_j\beta\beta'B_k\beta e_je_k+\frac{4}{n\phi^2}\beta'B\eta_0\sum_j\beta'B_j\beta e_j-\\&-\frac{2}{\phi\sqrt{n}}\beta'B\eta_{-\frac{1}{2}}-\frac{2}{\phi n}\sum_j\beta'B_j\eta_0e_j-\frac{1}{n\phi}\sum_{j,k}\beta'B_{jk}\beta e_je_k\Bigg)-\\&-\left(\frac{2}{\sqrt{n}}\beta'B\eta_0+\frac{1}{\sqrt{n}}\sum_j\beta'B_j\beta e_j\right)\left(\frac{2}{\phi\sqrt{n}}\beta'B\eta_0+\frac{1}{\phi\sqrt{n}}\sum_j\beta'B_j\beta e_j\right)+\\&+\frac{2}{n}\sum_j\beta'B_j\eta_0e_j+\frac{1}{n}\sum_{j,k}\beta'B_{jk}\beta e_je_k.\end{aligned}\tag{15}$$

Then we obtain the following results up to the order $O(n^{-1})$ *using the distributional properties of disturbances:*

$$E(\rho_{-\frac{1}{2}})=0\quad(\text{as }\ E(\eta_0)=0,E(e_j)=0\quad\text{using Lemma 1}),\tag{16}$$

$$\begin{aligned}E(\rho_{-1})=\beta'B\beta&\left[\frac{4}{n}\beta'B\beta+\frac{1}{n\sigma^2}\sum_{j,k}\beta'B_j\beta\beta'B_k\beta\lambda_{jk}-\frac{1}{n\phi}\sum_{j,k}\beta'B_{jk}\beta\lambda_{jk}\right]-\\&-\frac{4\sigma^2}{\phi n}\beta'B\beta-\frac{1}{\phi n}\sum_{j,k}\beta'B_j\beta\beta'B_k\beta\lambda_{jk}+\frac{1}{n}\sum_{j,k}\beta'B_{jk}\beta\lambda_{jk},\end{aligned}\tag{17}$$

$$\begin{aligned}E(\rho^2_{-\frac{1}{2}})=&\frac{4}{n}\sigma^2\beta'B\beta+\frac{1}{n}\sum_{j,k}\beta'B_j\beta\beta'B_k\beta\lambda_{jk}+\frac{1}{\phi^2n}(\beta'B\beta)^2\sum_{j,k}\beta'B_j\beta\beta'B_k\beta\lambda_{jk}-\\&-\frac{4}{\phi n}(\beta'B\beta)^2\sigma^2-\frac{2}{\phi n}\beta'B\beta\sum_{j,k}\beta'B_j\beta\beta'B_k\beta\lambda_{jk}.\end{aligned}\tag{18}$$

Theorem 1. *The* R^2_{fkk} *is expressible as*

$$R^2_{fkk}=\tilde{R}^2_{fkk}+O_p(n^{-3/2}),\tag{19}$$

where

$$\tilde{R}^2_{fkk} = \left(1 - \frac{2k_1\sigma^2}{n\xi}\right)\frac{1}{\phi}\Big[\beta' B\beta + \rho_{-\frac{1}{2}} + \rho_{-1}\Big]. \tag{20}$$

The first and second moments of $\tilde{R}^2_{fkk}$ are given as follows:

$$\begin{aligned} E(\tilde{R}^2_{fkk}) = \left(1 - \frac{2k_1\sigma^2}{n\xi}\right)\frac{1}{\phi}\Bigg[&\beta' B\beta - \frac{\beta' B\beta}{n}\sum_{j,k}\beta' B_{jk}\beta\lambda_{jk} + \frac{4\sigma^2}{n\phi^2}(\beta' B\beta)^2 + \\ &+ \frac{\beta' B\beta}{n\phi^2}\sum_{j,k}\beta' B_j\beta\beta' B_k\beta\lambda_{jk} - \frac{4\sigma^2}{n\phi}\beta' B\beta - \frac{1}{n\phi}\sum_{j,k}\beta' B_j\beta\beta' B_k\beta\lambda_{jk} + \\ &+ \frac{1}{n}\sum_{j,k}\beta' B_{jk}\beta\lambda_{jk} + \frac{\sigma^2}{n}\sum_{j,k}\mathrm{tr}(B^{-1}P_jAP'_k)\lambda_{jk}\Bigg] + O(n^{-3/2}), \end{aligned} \tag{21}$$

$$\begin{aligned} E(\tilde{R}^2_{fkk})^2 = \left(1 - \frac{2k_1\sigma^2}{n\xi}\right)^2\frac{1}{\phi^2}\Bigg[&\frac{8(\beta' B\beta)^3}{n} + \\ &+ (\beta' B\beta)^2\Bigg\{1 + \frac{\sum_{j,k}\beta' B_j\beta\beta' B_k\beta\lambda_{jk}}{n}\left(\frac{1}{\phi^2} + \frac{2}{\sigma^2}\right) - \\ &- \frac{2}{n\phi}\left(\sum_{j,k}\beta' B_{jk}\beta\lambda_{jk} + 6\sigma^2\right)\Bigg\} + \frac{\sum_{j,k}\beta' B_j\beta\beta' B_k\beta\lambda_{jk}}{n} + \\ &+ \frac{2(\beta' B\beta)}{n}\Bigg\{-\frac{2(\sum_{j,k}\beta' B_j\beta\beta' B_k\beta\lambda_{jk})}{\phi} + \sum_{j,k}\beta' B_{jk}\beta\lambda_{jk} + 2\sigma^2\Bigg\} + \\ &+ O(n^{-3/2}). \end{aligned} \tag{22}$$

The variance of $\tilde{R}^2_{fkk}$ can be obtained using (21) and (22) as

$$\mathrm{Var}(\tilde{R}^2_{fkk}) = E(\tilde{R}^2_{fkk})^2 - [E(\tilde{R}^2_{fkk})]^2.$$

Proof of Theorem 1. We can write $A\Omega^{-1}A = A$ and then

$$y'Ay = v + \hat{\beta}'_g X'AX\hat{\beta}_g.$$

We recall that $\vartheta = \frac{\beta'\Sigma_{XX}\beta}{\beta'\Sigma_{XX}\beta + \sigma^2}$; $0 \le \vartheta \le 1$, $\Sigma_{XX} = \mathrm{plim}_{n\to\infty}\frac{X'AX}{n}$ and $R^2_{fkk} = \frac{\hat{\beta}'_{kk}X'\hat{A}X\hat{\beta}_{kk}}{y'\hat{A}y}$, $0 \le R^2_{fkk} \le 1$. We write $A \equiv A\theta$, $\hat{A} \equiv A(\hat{\theta})$, where θ is a $q\times 1$ vector, and thus we can express

$$\hat{\beta}_{fg} = \beta + \frac{1}{\sqrt{n}}(\eta_0 + \eta_{-\frac{1}{2}}) + O_p(n^{-3/2}), \tag{23}$$

where $P_jX = 0$. Further, let

$$\hat{\delta}_{kk} = \left(\frac{k_1}{n-p+2}\right)\frac{\hat{v}}{\hat{\beta}'_{fg}X'\hat{A}X\hat{\beta}_{fg} + (1-k_2)\hat{v}}. \tag{24}$$

Consider the denominator of (24), we can write using (23)

$$\frac{1}{n}\Big[\hat{\beta}'_{fg}X'\hat{A}X\hat{\beta}_{fg} + (1-k_2)\hat{v}\Big] =$$

$$= \left(\beta + \frac{1}{\sqrt{n}}\eta_0 + \frac{1}{n}\eta_{-\frac{1}{2}}\right)'\left(B + \frac{1}{\sqrt{n}}\sum_j B_je_j\right)\left(\beta + \frac{1}{\sqrt{n}}\eta_0 + \frac{1}{n}\eta_{-\frac{1}{2}}\right) + (1-k_2)\frac{\hat{v}}{n} =$$

$$= \beta' B\beta + \frac{2}{\sqrt{n}}\beta' B\eta_0 + \frac{1}{\sqrt{n}}\sum_j \beta' B_j \beta e_j + (1-k_2)\sigma^2 + O_p(n^{-1}), \tag{25}$$

where η_0 and $\eta_{-\frac{1}{2}}$ are defined in Lemma 1.

Substituting (25) in (24), we get

$$\hat\delta_{kk} = \frac{k_1}{n} \quad \frac{\sigma^2}{\beta' B\beta + \frac{2}{\sqrt{n}}\beta' B\eta_0 + \frac{1}{\sqrt{n}}\sum_j \beta' B_j\beta e_j + (1-k_2)\sigma^2} =$$

$$= \frac{k_1}{n} \quad \frac{\sigma^2}{\beta' B\beta + (1-k_2)\sigma^2}\left[1 + \frac{2}{\xi\sqrt{n}}\beta' B\eta_0 + \frac{1}{\xi\sqrt{n}}\sum_j \beta' B_j\beta e_j\right]^{-1} =$$

$$= \frac{k_1}{n} \quad \frac{\sigma^2}{\xi}\left[1 - \frac{2}{\xi\sqrt{n}}\beta' B\eta_0 - \frac{1}{\xi\sqrt{n}}\sum_j \beta' B_j\beta e_j\right].$$

Next, we find

$$1-\hat\delta_{kk} = 1 - \frac{k_1\sigma^2}{n\xi} + \frac{k_1\sigma^2}{n\sqrt{n}\xi^2}\left(2\beta' B\eta_0 + \sum_j \beta' B_j\beta e_j\right),$$

$$(1-\hat\delta_{kk})^2 = 1 - \frac{2k_1\sigma^2}{n\xi} + \frac{2k_1\sigma^2}{n\sqrt{n}\xi^2}\left(2\beta' B\eta_0 + \sum_j \beta' B_j\beta e_j\right),$$

$$1-(1-\hat\delta_{kk})^2 = \frac{2k_1\sigma^2}{n\xi} - \frac{2k_1\sigma^2}{n\sqrt{n}\xi^2}\left(2\beta' B\eta_0 + \sum_j \beta' B_j\beta e_j\right).$$

Let $\hat B = \frac{1}{n}X'\hat A X$, $B_j = \frac{1}{n}X' A_j X$, $B_{jk} = \frac{1}{n}X' A_{jk} X$. We can express now

$$R^2_{fkk} = \frac{(1-\hat\delta_{kk})^2\hat\beta'_{fg}\left(\frac{1}{n}X'\hat A X\right)\hat\beta_{fg}}{\hat\beta'_{fg}\left(\frac{1}{n}X'\hat A X\right)\hat\beta_{fg} + \frac{1}{n}\hat v} = \frac{(1-\hat\delta_{kk})^2\hat\beta'_{fg}\hat B\hat\beta_{fg}}{\hat\beta'_{fg}\hat B\hat\beta_{fg} + \frac{1}{n}\hat v}. \tag{26}$$

Consider the numerator of (26)

$$\hat\beta'_{fg}\hat B_{fg}\hat\beta_{fg} = \left(\beta + \frac{1}{\sqrt{n}}\eta_0 + \frac{1}{\sqrt{n}}\eta_{-\frac{1}{2}}\right)\left(B + \frac{1}{\sqrt{n}}\sum_j B_j e_j + \frac{1}{n}\sum_{j,k} B_{jk} e_j e_k\right) \times$$

$$\times\left(\beta + \frac{1}{\sqrt{n}}\eta_0 + \frac{1}{\sqrt{n}}\eta_{-\frac{1}{2}}\right) =$$

$$= \beta' B\beta + \frac{2}{\sqrt{n}}\beta' B\eta_0 + \frac{1}{\sqrt{n}}\sum_j \beta' B_j\beta e_j + \frac{2}{\sqrt{n}}\beta' B\eta_{-\frac{1}{2}} +$$

$$+ \frac{2}{n}\sum_j \beta' B_j\beta\eta_0 e_j + \frac{1}{n}\sum_{j,k}\beta' B_{jk}\beta e_j e_k + \frac{1}{n}\eta'_{-\frac{1}{2}}B\eta_{-\frac{1}{2}} + O_p(n^{-3/2}). \tag{27}$$

Also $\frac{1}{n}\hat v = \sigma^2 + O_p(n^{-3/2})$ and $(1-\hat\delta_{kk})^2 = 1 - \frac{2k_1\sigma^2}{n\xi}$.

Consider the denominator of (26)

$$\frac{1}{\hat\beta'_{fg}\hat B\hat\beta_{fg} + \frac{1}{n}\hat v} = \Bigg[\beta' B\beta + \sigma^2 + \frac{2}{\sqrt{n}}\beta' B\eta_0 + \frac{1}{\sqrt{n}}\sum_j \beta' B_j\beta e_j + \frac{2}{\sqrt{n}}\beta' B\eta_{-\frac{1}{2}} +$$

$$+\frac{2}{n}\sum_j \beta' B_j \beta \eta_0 e_j + \frac{1}{n}\sum_{j,k} \beta' B_{jk} \beta e_j e_k + \frac{1}{n}\eta'_{-\frac{1}{2}} B \eta_{-\frac{1}{2}}\Bigg]^{-1} =$$

$$= \frac{1}{\phi}\Bigg[1 + \frac{2}{\phi\sqrt{n}}\beta' B\eta_0 + \frac{1}{\phi\sqrt{n}}\sum_j \beta' B_j \beta e_j + \frac{2}{\phi\sqrt{n}}\beta' B\eta_{-\frac{1}{2}} +$$

$$+\frac{2}{\phi n}\sum_j \beta' B_j \beta \eta_0 e_j + \frac{1}{\phi n}\sum_{j,k} \beta' B_{jk}\beta e_j e_k\Bigg]^{-1} =$$

$$= \frac{1}{\phi}\Bigg[1 - \frac{2}{\phi\sqrt{n}}\beta' B\eta_0 - \frac{1}{\phi\sqrt{n}}\sum_j \beta' B_j \beta e_j + \frac{4}{\phi^2 n}(\beta' B\eta_0)^2 +$$

$$+\frac{1}{\phi^2 n}\sum_{j,k} \beta' B_j \beta \beta' B_k \beta e_j e_k + \frac{4}{\phi^2 n}\beta' B\eta_0 \sum_j \beta' B_j \beta e_j -$$

$$-\frac{2}{\phi\sqrt{n}}\beta' B\eta_{-\frac{1}{2}} - \frac{2}{\phi n}\sum_j \beta' B_j \eta_0 e_j - \frac{1}{\phi n}\sum_{j,k}\beta' B_{jk}\beta e_j e_k\Bigg]. \tag{28}$$

Hence using (27) and (28) in (26), we expand R^2_{fkk} and obtain its expression as

$$R^2_{fkk} = \left(1 - \frac{2k_1\sigma^2}{n\xi}\right)\frac{1}{\phi}\Bigg[\beta' B\beta + \frac{2}{\sqrt{n}}\beta' B\eta_0 + \frac{1}{\sqrt{n}}\sum_j \beta' B_j\beta e_j + \frac{2}{\sqrt{n}}\beta' B\eta_{-\frac{1}{2}} +$$

$$+\frac{2}{n}\sum_j \beta' B_j \beta\eta_0 e_j + \frac{1}{n}\sum_{j,k}\beta' B_{jk}\beta e_j e_k + \frac{1}{n}\eta'_{-\frac{1}{2}} B\eta_{-\frac{1}{2}}\Bigg] \times \Bigg[1 - \frac{2}{\phi\sqrt{n}}\beta' B\eta_0 -$$

$$-\frac{1}{\phi\sqrt{n}}\sum_j \beta' B_j\beta e_j + \frac{4}{\phi^2 n}(\beta' B\eta_0)^2 + \frac{1}{\phi^2 n}\sum_{j,k}\beta' B_j\beta\beta' B_k\beta e_j e_k +$$

$$+\frac{4}{\phi^2 n}\beta' B\eta_0\sum_j \beta' B_j\beta e_j - \frac{2}{\phi\sqrt{n}}\beta' B\eta_{-\frac{1}{2}} - \frac{2}{\phi n}\sum_j \beta' B_j\eta_0 e_j -$$

$$-\frac{1}{\phi n}\sum_{j,k}\beta' B_{jk}\beta e_j e_k\Bigg] + O_p(n^{-3/2}) =$$

$$= \left(1 - \frac{2k_1\sigma^2}{n\xi}\right)\frac{1}{\phi}\Bigg[\beta' B\beta + \Bigg\{\frac{2}{\sqrt{n}}\beta' B\eta_0 + \frac{1}{\sqrt{n}}\sum_j \beta' B_j\beta e_j - \frac{\beta' B\beta}{\phi\sqrt{n}}\beta' B\eta_0 -$$

$$-\frac{1}{\phi\sqrt{n}}\beta' B\beta\sum_j \beta' B_j\beta e_j\Bigg\} + \Bigg\{\beta' B\beta\Bigg(\frac{4}{n\phi^2}(\beta' B\eta_0)^2 +$$

$$+\frac{1}{n\phi^2}\sum_{j,k}\beta' B_j\beta\beta' B_k\beta e_j e_k + \frac{4}{n\phi^2}\beta' B\eta_0\sum_j \beta' B_j\beta e_j -$$

$$-\frac{2}{\phi\sqrt{n}}\beta' B\eta_{-\frac{1}{2}} - \frac{2}{\phi n}\sum_j \beta' B_j\eta_0 e_j - \frac{1}{n\phi}\sum_{j,k}\beta' B_{jk}\beta e_j e_k\Bigg\} -$$

$$- \left(\frac{2}{\sqrt{n}} \beta' B \eta_0 + \frac{1}{\sqrt{n}} \sum_j \beta' B_j \beta e_j \right) \left(\frac{2}{\phi\sqrt{n}} \beta' B \eta_0 + \frac{1}{\phi\sqrt{n}} \sum_j \beta' B_j \beta e_j \right) +$$

$$+ \frac{2}{n} \sum_j \beta' B_j \eta_0 e_j + \frac{1}{n} \sum_{j,k} \beta' B_{jk} \beta e_j e_k \Bigg] + O_p(n^{-3/2}) =$$

$$= \left(1 - \frac{2k_1\sigma^2}{n\xi}\right) \frac{1}{\phi} \left[\beta' B \beta + \rho_{-\frac{1}{2}} + \rho_{-1} \right] + O_p(n^{-3/2}) =$$

$$= \tilde{R}^2_{fkk} + O_p(n^{-3/2}) , \qquad (29)$$

where $\tilde{R}^2_{fkk}$ is defined in (20); $\rho_{-\frac{1}{2}}$ and ρ_{-1} are defined in Lemma 2. This is the result (19) reported in Theorem 1.

Thus, using the results in Lemma 1, we obtain the expectation of R^2_{fkk} as

$$E(\tilde{R}^2_{fkk}) = \left(1 - \frac{2k_1\sigma^2}{n\xi}\right) \frac{1}{\phi} \Bigg[\beta' B \beta - \frac{\beta' B \beta}{n} \sum_{j,k} \beta' B_{jk} \beta \lambda_{jk} + \frac{4\sigma^2}{n\phi^2} (\beta' B \beta)^2 +$$

$$+ \frac{\beta' B \beta}{n\phi^2} \sum_{j,k} \beta' B_j \beta \beta' B_k \beta \lambda_{jk} - \frac{4\sigma^2}{n\phi} \beta' B \beta -$$

$$- \frac{1}{n\phi} \sum_{j,k} \beta' B_j \beta \beta' B_k \beta \lambda_{jk} + \frac{1}{n} \sum_{j,k} \beta' B_{jk} \beta \lambda_{jk} +$$

$$+ \frac{\sigma^2}{n} \sum_{j,k} \mathrm{tr}(B^{-1} P_j A P_k') \lambda_{jk} \Bigg] + O(n^{-3/2}),$$

which is the result (21) reported in Theorem 1.

The second moment of R^2_{fkk} up to the order $O(n^{-1})$ is obtained as

$$E(\tilde{R}^2_{fkk})^2 = \left(1 - \frac{2k_1\sigma^2}{n\xi}\right)^2 \frac{1}{\phi^2} E\left[\beta' B \beta + \rho_{-\frac{1}{2}} + \rho_{-1} \right]^2 + O(n^{-3/2}) =$$

$$= \left(1 - \frac{2k_1\sigma^2}{n\xi}\right)^2 \frac{1}{\phi^2} \left[(\beta' B \beta)^2 + 2\beta' B \beta E(\rho_{-\frac{1}{2}}) + E(\rho^2_{-\frac{1}{2}}) + 2\beta' B \beta E(\rho_{-1}) \right] +$$

$$+ O(n^{-3/2}). \qquad (30)$$

Substituting the expressions of $E(\rho_{-\frac{1}{2}})$, $E(\rho_{-1})$ and $E(\rho^2_{-\frac{1}{2}})$ using Lemma 2 in (30), we obtain

$$E(\tilde{R}^2_{fkk})^2 = \left(1 - \frac{2k_1\sigma^2}{n\xi}\right)^2 \frac{1}{\phi^2} \Bigg[\frac{8(\beta' B \beta)^3}{n} +$$

$$+ (\beta' B \beta)^2 \Bigg\{ 1 + \frac{\sum_{j,k} \beta' B_j \beta \beta' B_k \beta \lambda_{jk}}{n} \left(\frac{1}{\phi^2} + \frac{2}{\sigma^2} \right) -$$

$$- \frac{2}{n\phi} \left(\sum_{j,k} \beta' B_{jk} \beta \lambda_{jk} + 6\sigma^2 \right) \Bigg\} + \frac{\sum_{j,k} \beta' B_j \beta \beta' B_k \beta \lambda_{jk}}{n} +$$

$$+ \frac{2(\beta' B \beta)}{n} \left\{ - \frac{2(\sum_{j,k} \beta' B_j \beta \beta' B_k \beta \lambda_{jk})}{\phi} + \sum_{j,k} \beta' B_{jk} \beta \lambda_{jk} + 2\sigma^2 \right\} +$$

$$+ O(n^{-3/2}),$$

which is the result (22) reported in Theorem 1.

The variance of $\tilde{R}^2_{fkk}$ can be obtained using (21) and (22) as $\mathrm{Var}(\tilde{R}^2_{fkk}) = E(\tilde{R}^2_{fkk})^2 - [E(\tilde{R}^2_{fkk})]^2$. This completes the proof of Theorem 1. □

The results stated in Theorem 1 are useful from several perspectives. It may be noted that all the statistics which are proposed to measure the goodness of fit are essentially a function of random variables, i.e., a statistic and they are aiming to consistently estimate the population correlation coefficient between the study variable and explanatory variables. The objective of any statistician is not only to estimate the parameters but also to use it for other statistical procedures like test of hypothesis and confidence intervals of the parameter etc. Obviously, any statistician would like to use the exact moments for such purposes but since the exact moments are difficult to derive, the approximate moments derived in this paper will help. So the moments of GoFs derived will help in deriving the statistical inference about the population correlation coefficient between the study variable and explanatory variables.

5. Simulation study

The approximate moments of the DKKE and FDKKE up to the first order of approximation essentially describe the behaviour of the estimators when sample size is large. To get an idea about the performance of the proposed goodness-of-fit statistics as well as their feasible versions in finite samples, we conducted a detailed Monte-Carlo simulation study. To save the space, we are presenting here few results only. The detailed results are available at `http://home.iitk.ac.in/~shalab/r2dkcsc.pdf`.

To understand the effect of non-identity structure of covariance matrix, we consider the random errors ϵ in (7) to follow an autoregressive process AR(1) as $\epsilon_i = \rho\epsilon_{i-1} + v_i$ where $|\rho| < 1$ is the autocorrelation coefficient between ϵ_i and ϵ_{i-1}, $(i = 1, 2, \ldots, n)$; and v_i's are identically and independently distributed $N(0, \sigma_v^2)$. We assume $\sigma^2 = 1$ and so the covariance matrix of ϵ is given by

$$V(\epsilon) = \Omega^{-1} = \frac{1}{1-\rho^2}\begin{pmatrix} 1 & \rho & \ldots & \rho^{n-1} \\ \rho & 1 & \ldots & \rho^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \ldots & 1 \end{pmatrix}$$

and consequently

$$\Omega = \begin{pmatrix} 1 & -\rho & 0 & \ldots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \ldots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \ldots & -\rho & 1 \end{pmatrix}.$$

The autocorrelation coefficient ρ can be estimated by $\hat{\rho} = \sum_{i=2}^{n-1} \hat{\epsilon}_i\hat{\epsilon}_{i-1} / \sum_{i=2}^{n} \hat{\epsilon}_{i-1}^2$, where $\hat{\epsilon}_i = y_i - x_i'\hat{\beta}_{kk} (i = 1, 2, \ldots, n)$ are the residuals based on $\hat{\beta}_{kk}$. When using FDKKE, the $\hat{\epsilon}_i$'s are obtained by using $\hat{\beta}_{fkk}$ in place of $\hat{\beta}_{kk}$.

We consider the following versions of DKKE and FDKKE that are obtained by substituting the values of k_1 and k_2. When $k_1 = 0$, we get GLSE and FGLSE. When $k_1 = p - 2$ and $k_2 = 1$, we get generalized Stein-rule estimator (GSRE) and feasible generalized Stein-rule estimator (FGSRE). When $k_1 = \frac{1}{n-p}$ and $k_2 = 1 - k_1$, we get generalized minimum mean squared error estimator (GMMSE) and feasible generalized minimum mean squared error estimator (FGMMSE), proposed by [16], see also [10, 11].

When $k_1 = \frac{n-p+2}{n-p}$ and $k_2 = 1 - \frac{k_1}{n-p+2}$, we get adjusted generalized minimum mean squared error estimator (AGMMSE) and adjusted feasible generalized minimum mean squared error estimator (AFGMMSE), proposed by [11]. When $k_1 = \frac{(n-p+2)p}{n-p}$ and $k_2 = 1 - \frac{k_1}{n-p+2}$, we get generalized double k-class estimators (GKKCE) and feasible generalized double k-class estimators (FGKKCE), proposed by [15].

We considered the cases when $n = 20$, $n = 35$ and $n = 100$ that indicate the small, moderately large and large sample sizes respectively. The results beyond $n = 100$ generally stabilize and change very little as n increases further. The traditional R^2 (coefficient of determination based on OLSE) increases as the number of explanatory variables increases and we also want to check if such behavior continues to hold in our proposed goodness-of-fit statistics. We have chosen $p = 4$ and $p = 8$ to see the effect of increased number of explanatory variables on various goodness-of-fit statistics. To understand the effect of σ_v^2, we consider $\sigma_v^2 = 1$, $\sigma_v^2 = 5$ and $\sigma_v^2 = 10$ that are considered to represent the lower, moderately large and higher variances respectively in the data. Based on these choices, we have considered various combinations of these values. The data is generated on ϵ's using AR(1) process with $\rho = -0.9, -0.8, \ldots, -0.1, 0, 0.1, \ldots, 0.8, 0.9$. Then the values of study variable are generated for each ρ. The values of explanatory variables are held fixed for all the values of given autocorrelation coefficients so that the values of the estimated regression coefficients and in turn, all proposed R^2_{kk} and R^2_{fkk} statistics are computed for the same set of data on X and y. This will enable us to compare the versions of R^2_{kk} statistics as they will be affected by the same sources and magnitude of variability. First, we estimate β by assuming Ω^{-1} to be known and obtain the simulated values of goodness-of-fit statistic R^2_{kk}. Based on these values, we find the empirical absolute relative bias (RB) and empirical relative mean squared error (RM) of all the goodness-of-fit statistics under consideration. We execute this process for a given value of ρ. In the next step, we estimate the same ρ as $\hat{\rho}$ on the basis of the the data set generated on X and y in the earlier step. Using thus obtained $\hat{\rho}$, we obtain $\hat{\Omega}$ and then further obtain the simulated values of goodness-of-fit statistic R^2_{fkk}. So the X and y values used in computing R^2_{kk} and R^2_{fkk} are the same and this helps in studying the changes occurring between the estimators and their feasible versions. Again, based on these values, we find the RB and RM of R^2_{fkk}. The RB and RM of all the R^2_{kk} and R^2_{fkk} statistics are computed based on 5000 replications for various combinations of n, p and σ_v^2. Some representative selected values of RB and RM of proposed statistics are presented in Tables 1–2. We also have constructed the three dimensional surface plots of RBs and RMs with respect to ρ and ϑ to further investigate their joint effect on the behavior of proposed statistics, see e. g., Figures 1–2. We have presented only representative and selected number of tables and figures. More description about tables and figures is available at `http://home.iitk.ac.in/~shalab/r2dkcsc.pdf`.

The empirical values of relative absolute bias (RB) and relative mean squared errors (RM) of an estimator $\hat{\delta}$ of a parameter δ based on G replications are calculated as $RB(\hat{\delta}) = \frac{1}{g}\sum_{g=1}^{G}\left|\frac{(\hat{\delta}_g - \delta)}{\delta}\right|$ and $RM(\hat{\delta}) = \frac{1}{g}\sum_{g=1}^{G}\frac{\left(\hat{\delta}_g - \delta\right)^2}{\delta^2} + [RB(\hat{\delta})]^2$ respectively, where $\hat{\delta}_g$ is based on a sample of size n. Here, the notations R^2_{gls}, R^2_{fgls}, R^2_{gsre}, R^2_{fgsre}, R^2_{gmmse}, R^2_{fgmmse}, R^2_{agmmse}, $R^2_{fagmmse}$, R^2_{gkkc} and R^2_{fgkkce} denote the goodness-of-fit statistic based on GLSE, FGLSE, GSRE, FGSRE, GMMSE, FGMMSE, AGMMSE, FAGMMSE, GKCCE and GKCCFE respectively. For example, $R^2_{gls} = \frac{\hat{\beta}'_{gls}X'AX\hat{\beta}_{gls}}{y'Ay}$ and $R^2_{fgls} = = \frac{\hat{\beta}'_{fgls}X'AX\hat{\beta}_{fgls}}{y'Ay}$, where $\hat{\beta}_{gls} = (X'AX)^{-1}X'Ay$ and $\hat{\beta}_{fgls} = (X'\hat{A}X)^{-1}X'\hat{A}y$.

One may note that $\mathrm{Var}(\epsilon) = \sigma_v^2/(1-\rho^2)$, so whenever ρ is close to ± 1, the variance is expected to be very high and it is not advisable to have such a high variance σ_v^2, otherwise the data will contain enormous variation. This fact is reflected in all the tables showing that the values of σ_ϵ^2 decrease as the values of $|\rho|$ decrease.

First, we consider the effect of increasing sample size on the behavior of different estimators when $p = 4$ (small) and $\sigma_v^2 = 1$ (low) remains the same and only n increases from $n = 20, 35$ to $n = 100$. A common feature is that the RB of all R_{kk}^2 and R_{fkk}^2 statistics decrease as n increases. The same conclusion continues to hold true when the values of $|\rho|$ decrease for fixed n, p and σ_v^2. The behavior of RBs is nearly symmetric for positive and negative values of ρ and only slight variation is present which is due to the variation in the values of random errors v's. The behavior of RBs of R_{kk}^2 and R_{fkk}^2 statistics when $|\rho|$ is high, say 0.8 and 0.9 remains stable provided σ_v^2 is smaller (here $\sigma_v^2 = 1$). The effect of $|\rho|$ on R_{kk}^2 and R_{fkk}^2 statistics becomes more prominent when σ_v^2 increases (say $\sigma_v^2 = 10$) but this effect diminishes if the sample size becomes higher. An important observation emerging from simulated values is that only $R_{fagmmse}^2$, R_{fgsre}^2, R_{gkkc}^2 and R_{fgkkc}^2 show abrupt behaviors for $|\rho| = 0.9$ when n and p are not very large (say $n = 20, p = 4$) and σ_v^2 is large ($\sigma_v^2 = 10$). As soon as either n or p becomes larger, then the behaviours of $R_{fagmmse}^2$, R_{fgsre}^2, R_{gkkc}^2 and R_{fgkkc}^2 improve. This is not a surprising outcome because the data in such a case also has large variation which is being controlled by the total number of observations. All other R_{kk}^2 and R_{fkk}^2 statistics behave well in all the situations even when $|\rho| = 0.9$. So the proposed R_{kk}^2 and R_{fkk}^2 statistics can also be used in case the data has near unit root problem. In general, R_{gsre}^2, R_{fgsre}^2, R_{gkkc}^2 and R_{fgkkc}^2 are not strongly recommended for near unit root situations. When the number of explanatory variables becomes higher, R_{kk}^2 and R_{fkk}^2 statistics perform well even for small sample size and large σ_v^2 values. In a nutshell, under the criterion of RBs, the R_{kk}^2 and R_{fkk}^2 statistics works well when σ_v^2 is small (irrespective of n), σ_v^2 and n both are large or σ_v^2 is large, n is small but p is large.

As a rule of thumb, it is indicated that the effect of σ_v^2 is controlled by the size of X matrix which contains np observations. If np is large, it can subside the effect of large error variation.

Another important strength of the proposed R_{kk}^2 and R_{fkk}^2 statistics is that they are not much dependent on the values of ϑ. The RBs of all the R_{kk}^2 and R_{fkk}^2 statistics are usually small irrespective of the values of ϑ. We have simulated the data in which ϑ is as low as 0.3 and as high as 0.95. The RBs of all the R_{kk}^2 and R_{fkk}^2 statistics for all the estimates are not usually high in most of the situations.

A good feature of the proposed goodness-of-fit statistics is that there is not much difference between the values of R_{kk}^2 and R_{fkk}^2 statistics. It has an important implication that in most of the cases in real data situations, the values of parameters are unknown. Usually, they are estimated on the basis of given a sample of data and plugged in the estimator. The use of proposed goodness-of-fit statistics and such analysis gives an idea about the magnitude of change in the fitting of model.

When σ_v^2 is low, then the general ordering (may not necessarily always hold true) of R_{kk}^2 and R_{fkk}^2 statistics emerging from the empirical results are

$$RB(R_{gmmse}^2) < RB(R_{gls}^2) < RB(R_{agmmse}^2) < RB(R_{gsre}^2) < RB(R_{gkkc}^2) \text{ and}$$

$$RB(R_{fgmmse}^2) < RB(R_{fgls}^2) < RB(R_{fagmmse}^2) < RB(R_{fgsre}^2) < RB(R_{fgkkc}^2).$$

Now, we consider the relative mean squared errors (RMs) of various R_{kk}^2 and R_{fkk}^2 statistics under different estimators. The RMs of all the R_{kk}^2 and R_{fkk}^2 statistics decreases as n increases, particularly when p and σ_v^2 stay fixed. The RMs of R_{gls}^2, R_{fgls}^2, R_{gmmse}^2 and R_{fgmmse}^2 are not significantly affected when $|\rho|$ is very high, say $|\rho| \geq 0.8$ but all other R_{kk}^2 and R_{fkk}^2 values based on AGMMSE, GSRE and GKCCE are seriously affected even when n and p are large enough. When both n and p increase but σ_v^2 decreases, the performances of these statistics remain unsatisfactory when $|\rho| \geq 0.8$. So GLSE and GMMSE along with their feasible versions provide a good fitted model even in the presence of a near unit root problem in the data. The RMs of all R^2_{kk} and R^2_{fkk} statistics usually decrease when p increases and/or σ_v^2 decreases. The performance is better for lower values of σ_v^2. The larger values of n and p are controlling the bad effects arising due to large values of σ_v^2 on the RMs of R^2_{kk} and R^2_{fkk}. A general rule of thumb emerging from the empirical values is that the performance of R^2_{kk} and R^2_{fkk} is better when np is large provided σ_v^2 is not too high. The values of RMs of all R^2_{kk} and R^2_{fkk} statistics decrease as $|\rho|$ decreases for a given n, p and σ_v^2. A general pattern of ordering of RMs of R^2_{kk} and R^2_{fkk} values emerging from the simulated results are

$$RM(R^2_{gls}) < RM(R^2_{gmmse}) < RM(R^2_{agmmse}) < RM(R^2_{gsre}) < RM(R^2_{gkkc}) \text{ and}$$

$$RM(R^2_{fgls}) < RM(R^2_{fgmmse}) < RM(R^2_{fagmmse}) < RM(R^2_{fgsre}) < RM(R^2_{fgkkc}).$$

The lowest and highest values of ϑ for which the data is simulated are close to 0.3 and 0.95 respectively. So the proposed goodness-of-fit statistics perform well in terms of RMs even for lower values of ϑ.

Under the criterion of RM, the proposed R^2_{kk} and R^2_{fkk} statistics can be used and will provide good results when both n and σ_v^2 are small, n is large and σ_v^2 is not too large or n and/or p are large such that np is large and σ_v^2 is not too large.

A notable observation from the results is that the dominance of an estimator does not guarantee a good fitted model based on the same estimator in terms of RM. For example, the Stein rule estimator is well known to provide more efficient estimator of regression coefficients than the least squares estimators like OLSE (or equivalently GLSE) when $p > 2$. The simulation results are indicating that even when $p > 2$, the use of Stein rule estimator does not guarantee a good fitted model in terms of COD. So the efficiency property of estimators does not necessarily guarantee the best fitted model too. This exemplifies the importance of the proposed goodness-of-fit statistics in this paper.

Another notable feature emerging from the simulated results is that there is not much difference between the R^2_{kk} and R^2_{fkk} statistics based on an estimator and its feasible version. In practice, the values of parameters are unknown and usually they are replaced by their estimated counterparts. Such a result ensures that the proposed goodness-of-fit statistics based on feasible various are nearly as good as the estimators themselves and there is not much change in the model performance in terms of RMs of the statistics.

The role and effect of positive and negative values of ρ is almost the same on RMs of all the R^2_{kk} and R^2_{fkk} statistics. The difference between the RMs at $+\rho$ and at $-\rho$ is minor provided $|\rho| < 0.8$.

The usual COD based on OLSE statistic in the classical linear regression model increases when p increases. We also compared the values of proposed R^2_{kk} and R^2_{fkk} by changing $p = 4$ to $p = 8$ and keeping all other parameters fixed as $n = 100$, $\sigma_v^2 = 5$. The results about RB and RM of R^2_{kk} and R^2_{fkk} based on all the estimators are presented in Tables 1-2.

It is clearly evident that the RB and RM for the case when $p = 8$ are smaller than in the case when $p = 4$. Hence R^2_{kk} and R^2_{fkk} increase as p increases.

To get more insight into the behaviors of R^2_{kk} and R^2_{fkk}, we have plotted the three-dimensional surface plots of RBs and RMs of R^2_{kk} and R^2_{fkk} versus various values of autocorrelation coefficient ρ and multiple correlation coefficient ϑ. This will give us an idea about the simultaneous effect of ρ and ϑ over R^2_{kk} and R^2_{fkk}. The plots in figures are arranged in the order of decreasing RBs and RMs from up to down. The graphs of RBs are arranged in the order as $RB(R^2_{gmmse})$, $RB(R^2_{gls})$, $RB(R^2_{agmmse})$, $RB(R^2_{gsre})$ and

$RB(R^2_{gkkc})$ from up to down. Similarly, the plots of RMs are arranged in the decreasing order as $RM(R^2_{gls})$, $RM(R^2_{gmmse})$, $RM(R^2_{agmmse})$, $RM(R^2_{gsre})$ and $RM(R^2_{gkkc})$ from up to down.

The surface plots for some selected values of n, p and σ_v^2 are presented in Figures 1–2 with the following notations: ρ and ϑ are denoted by 'rho' and 'theta' respectively; R^2_{gls}, R^2_{gmmse}, R^2_{agmmse}, R^2_{gsre} and R^2_{gkkc} are denoted as 'R2gls', 'R2gmmse', 'R2agmmse', 'R2gsre' and 'R2gkkc' respectively. Similarly R^2_{fgls}, R^2_{fgmmse}, $R^2_{fagmmse}$, R^2_{fgsre} and R^2_{fgkkc} are denoted as 'R2fgls', 'R2fgmmse', 'R2fagmmse', 'R2fgsre' and 'R2gkkc' respectively. The left and right panels of all the figures contain the RBs (or RMs) of the estimators and their corresponding feasible versions, respectively. So the side by side comparison of two figures gives a clear comparison of R^2_{kk} and R^2_{fkk} i.e., the estimator and its corresponding feasible versions. In some cases, the values of ρ are close to ± 0.9, then RBs the RMs are very high and they are recomputed after deleting the extreme values.

Now, we consider the three-dimensional surface plots from the figures about RBs. First, we consider the effect of n over the RBs of different statistics from Figures 1 and 2, where n increases from 20 to 100 keeping $\sigma_v^2 = 1$ and $p = 4$ fixed. The structure, curvature and pattern of all the figures based on the estimators and their feasible versions are nearly the same with an exception of R^2_{fgls} and R^2_{fgsr}. The RBs of R^2_{fgls} and R^2_{fgsr} gets stabilized only when n is large. The perturbations on surface decrease and smoothness increases as sample size increases. An increase in the value of σ_v^2 increases the perturbations on the surfaces. The values of σ_v^2 have stronger impact on the values of RB than n and p. Once the structure of surfaces stabilizes, the slope of the surfaces is decreasing as ϑ increases and/or $|\rho|$ decreases which indicates that the RBs are getting lower. The RBs seems to be more sensitive to the values of σ_v^2 than the values of n and p. Higher the np, more smooth is the surface provided σ_v^2 is not very large. The slopes of all the surfaces of RBs in all the figures are decreasing as ρ and ϑ decrease which indicates that the RBs decrease.

Now, we study the behaviour of RMs of the R^2_{kk} and R^2_{fkk} statistics for various combination of n, p and σ_v^2 from Figures. It is clear from these figures that when n increases, keeping p and σ_v^2 fixed but letting σ_v^2 be lower ($\sigma_v^2 = 1$), the structure, smoothness, pattern and perturbations on the surface of the plots gets stabilized. The increasing values of σ_v^2 tends to change the structure of the surfaces more dominantly than the values of n and p. If σ_v^2 increases, the side effects are controlled by np up to certain extent only. The surface plots of R^2_{gsre} and R^2_{gkkc} are more affected with σ_v^2 than other plots in most of the cases.

The slopes of surface and involved perturbations indicates the utility of proposed R^2_{kk} and R^2_{fkk} statistics with respect to ϑ. It is clear that when ϑ is low, the RMs are higher. As ϑ increases, the slope of the surfaces also decreases indicating that the RMs are getting lower. Again, this is expected because the lower values of ϑ indicate that either the data has very high variability beyond the ability of R^2_{kk} and R^2_{fkk} to handle or the underlying model that is being followed by the data is nonlinear. Such situations are correctly captured by the proposed R^2_{kk} and R^2_{fkk} statistics.

By comparing the surface plots on the left and right panels of the figures, one can conclude about how the performance of an estimator change when it is converted into a feasible form. We observe that all the figures on the left and right panels of the figures are not always similar. The degree of similarity of figures is generally high when np is large and σ_v^2 is low. This clearly indicates that the proposed R^2_{kk} and R^2_{fkk} statistics can be used well even when the unknown parameters are estimated and replaced by their estimated values depending upon the values of np and σ_v^2.

Table 1. Relative bias (RB) of R^2's when $\sigma_v^2 = 1$, $n = 100$ and $p = 4$

ρ	ϑ	σ_ϵ^2	$RB(R^2_{gls})$	$RB(R^2_{fgls})$	$RB(R^2_{gsre})$	$RB(R^2_{fgsre})$	$RB(R^2_{gmmse})$
-0.9	0.854	5.263	0.0847	0.0914	0.8053	0.8344	0.09
-0.8	0.91	2.777	0.0422	0.0455	0.5202	0.5372	0.0451
-0.7	0.928	1.960	0.025	0.0273	0.3883	0.3995	0.0271
-0.6	0.937	1.562	0.0153	0.0169	0.3163	0.3237	0.017
-0.5	0.941	1.333	0.0092	0.0101	0.274	0.2779	0.0107
-0.4	0.943	1.190	0.0044	0.0048	0.2448	0.2466	0.0057
-0.3	0.944	1.098	0.0012	0.0011	0.2272	0.2264	0.0024
-0.2	0.944	1.041	0.0012	0.0016	0.2156	0.2131	0.0018
-0.1	0.944	1.010	0.0024	0.0032	0.2102	0.2058	0.0013
0	0.944	1	0.0029	0.004	0.2084	0.2026	0.0018
0.1	0.944	1.010	0.0022	0.0035	0.2123	0.2054	0.001
0.2	0.943	1.041	0.001	0.0024	0.2187	0.2114	0.0002
0.3	0.943	1.098	0.0018	0.0003	0.2331	0.2256	0.0031
0.4	0.942	1.190	0.0056	0.004	0.2543	0.2466	0.007
0.5	0.940	1.333	0.0105	0.009	0.2846	0.2773	0.0121
0.6	0.935	1.562	0.0171	0.0156	0.3296	0.3232	0.0189
0.7	0.926	1.960	0.0276	0.0258	0.4052	0.3991	0.0298
0.8	0.907	2.777	0.0457	0.0437	0.5403	0.5371	0.0487
0.9	0.851	5.263	0.0911	0.0882	0.8273	0.8331	0.0967

ρ	ϑ	σ_ϵ^2	$RB(R^2_{fgmmse})$	$RB(R^2_{agmmse})$	$RB(R^2_{fagmmse})$	$RB(R^2_{gkkc})$	$RB(R^2_{fgkkc})$
-0.9	0.854	5.263	0.0967	0.5258	0.5454	0.929	0.9225
-0.8	0.91	2.777	0.0485	0.3074	0.3161	0.8313	0.8572
-0.7	0.928	1.960	0.0294	0.2213	0.2267	0.6708	0.6931
-0.6	0.937	1.562	0.0186	0.1759	0.1793	0.5677	0.5852
-0.5	0.941	1.333	0.0116	0.1496	0.1509	0.5032	0.5151
-0.4	0.943	1.190	0.0062	0.1313	0.1315	0.4577	0.4659
-0.3	0.944	1.098	0.0024	0.1202	0.1189	0.4299	0.4337
-0.2	0.944	1.041	0.0004	0.1128	0.1104	0.4116	0.4124
-0.1	0.944	1.010	0.0021	0.1093	0.1057	0.4032	0.4008
0	0.944	1	0.0029	0.1081	0.1037	0.4004	0.3957
0.1	0.944	1.010	0.0024	0.1105	0.1053	0.4067	0.4003
0.2	0.943	1.041	0.0012	0.1146	0.1091	0.417	0.4099
0.3	0.943	1.098	0.0015	0.1237	0.118	0.4397	0.4327
0.4	0.942	1.190	0.0053	0.1371	0.131	0.473	0.4663
0.5	0.940	1.333	0.0104	0.156	0.1501	0.5198	0.5148
0.6	0.935	1.562	0.0173	0.1842	0.1785	0.5875	0.5845
0.7	0.926	1.960	0.0279	0.2322	0.2258	0.6935	0.6928
0.8	0.907	2.777	0.0466	0.3213	0.3153	0.8519	0.8573
0.9	0.851	5.263	0.0936	0.5474	0.544	0.9078	0.9217

Table 2. Relative mean squared error (RM) of R^2's when $\sigma_v^2 = 1$, $n = 100$ and $p = 4$

ρ	ϑ	σ_ϵ^2	$RM(R^2_{gls})$	$RM(R^2_{fgls})$	$RM(R^2_{gsre})$	$RM(R^2_{fgsre})$	$RM(R^2_{gmmse})$
-0.9	0.8548	5.2632	0.002	0.0098	0.6595	0.7034	0.0102
-0.8	0.91	2.7778	0.0006	0.0025	0.2789	0.2947	0.0027
-0.7	0.9286	1.9608	0.0003	0.001	0.156	0.1636	0.0011
-0.6	0.9371	1.5625	0.0002	0.0004	0.1036	0.1075	0.0005
-0.5	0.9413	1.3333	0.0001	0.0002	0.0779	0.0795	0.0003
-0.4	0.9434	1.1905	0.0001	0.0001	0.0621	0.0627	0.0001
-0.3	0.9442	1.0989	0.0001	0.0001	0.0534	0.0529	0.0001
-0.2	0.9445	1.0417	0.0001	0.0001	0.048	0.0469	0.0001
-0.1	0.9444	1.0101	0.0001	0.0001	0.0456	0.0438	0.0001
0	0.9443	1	0.0001	0.0001	0.0448	0.0425	0.0001
0.1	0.9441	1.0101	0.0001	0.0001	0.0465	0.0436	0.0001
0.2	0.9439	1.0417	0.0001	0.0001	0.0494	0.0462	0.0001
0.3	0.9434	1.0989	0.0001	0.0001	0.0562	0.0525	0.0001
0.4	0.9423	1.1905	0.0001	0.0001	0.067	0.0627	0.0002
0.5	0.9401	1.3333	0.0001	0.0002	0.0838	0.079	0.0003
0.6	0.9357	1.5625	0.0002	0.0004	0.1124	0.1072	0.0006
0.7	0.9269	1.9608	0.0003	0.0009	0.1696	0.1633	0.0012
0.8	0.9078	2.7778	0.0006	0.0023	0.3002	0.2945	0.003
0.9	0.8514	5.2632	0.0021	0.0092	0.6947	0.7016	0.0115

ρ	ϑ	σ_ϵ^2	$RM(R^2_{fgmmse})$	$RM(R^2_{agmmse})$	$RM(R^2_{fagmmse})$	$RM(R^2_{gkkc})$	$RM(R^2_{fgkkc})$
-0.9	0.8548	5.2632	0.0108	0.2873	0.3052	0.8762	0.8633
-0.8	0.91	2.7778	0.0028	0.0989	0.1032	0.6996	0.7405
-0.7	0.9286	1.9608	0.0011	0.0514	0.0533	0.459	0.487
-0.6	0.9371	1.5625	0.0005	0.0325	0.0334	0.3297	0.348
-0.5	0.9413	1.3333	0.0003	0.0236	0.0237	0.2595	0.2704
-0.4	0.9434	1.1905	0.0001	0.0182	0.0181	0.2146	0.2216
-0.3	0.9442	1.0989	0.0001	0.0152	0.0148	0.1891	0.1921
-0.2	0.9445	1.0417	0.0001	0.0134	0.0128	0.1733	0.1738
-0.1	0.9444	1.0101	0.0001	0.0125	0.0118	0.1661	0.1643
0	0.9443	1	0.0001	0.0123	0.0113	0.1639	0.1602
0.1	0.9441	1.0101	0.0001	0.0128	0.0117	0.169	0.1639
0.2	0.9439	1.0417	0.0001	0.0138	0.0125	0.1778	0.1718
0.3	0.9434	1.0989	0.0001	0.0161	0.0146	0.1977	0.1911
0.4	0.9423	1.1905	0.0001	0.0198	0.018	0.2291	0.2219
0.5	0.9401	1.3333	0.0002	0.0256	0.0234	0.2764	0.2698
0.6	0.9357	1.5625	0.0005	0.0356	0.0331	0.3527	0.3473
0.7	0.9269	1.9608	0.001	0.0565	0.0529	0.4899	0.4867
0.8	0.9078	2.7778	0.0026	0.1078	0.1027	0.7335	0.7405
0.9	0.8514	5.2632	0.0103	0.3108	0.304	0.8456	0.8627

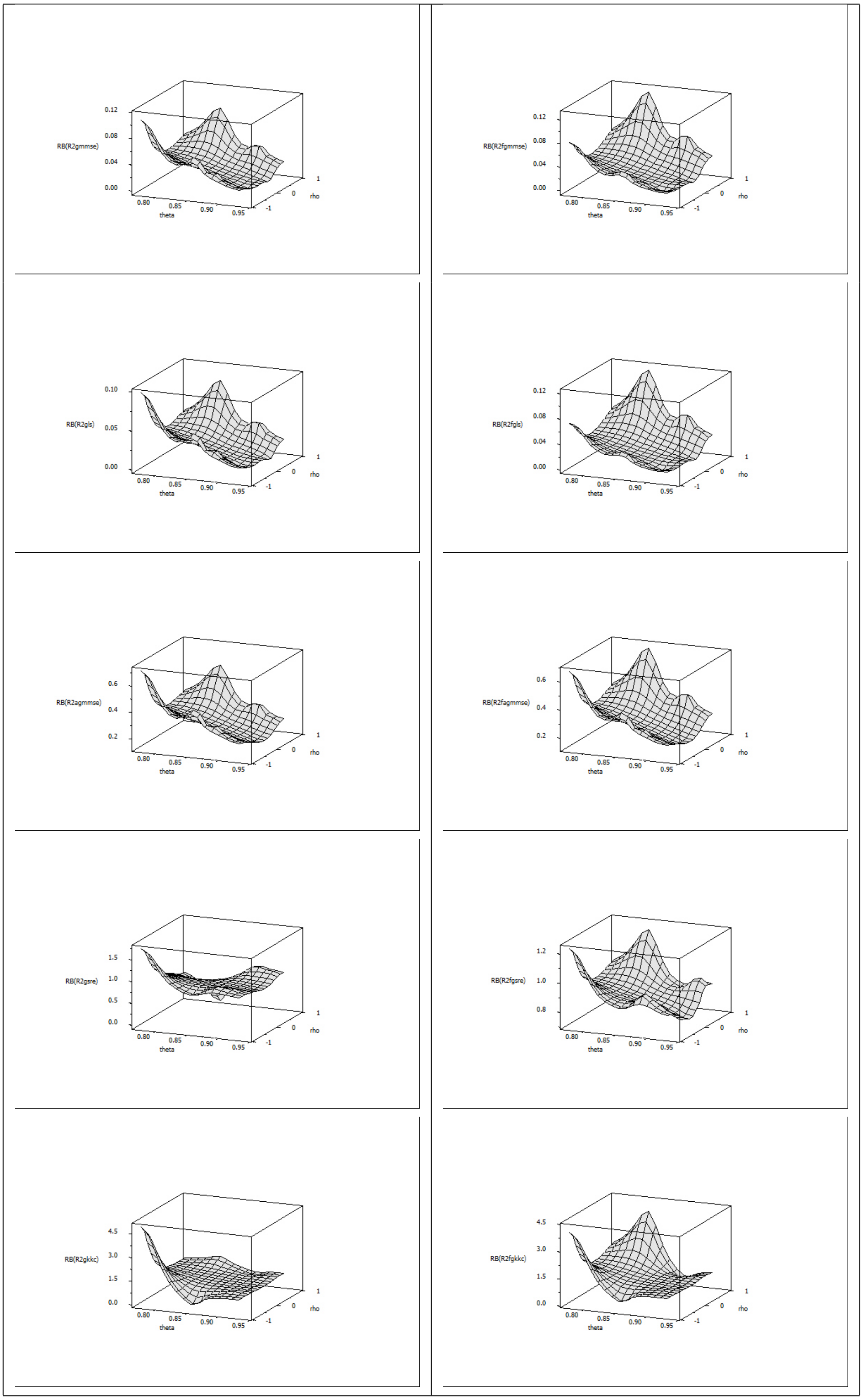

FIGURE 1. Surface plots of relative bias (RB) of R^2_{kk} and $R^2_{fkk}$when $\sigma_v^2 = 10$, $n = 100$ and $p = 8$

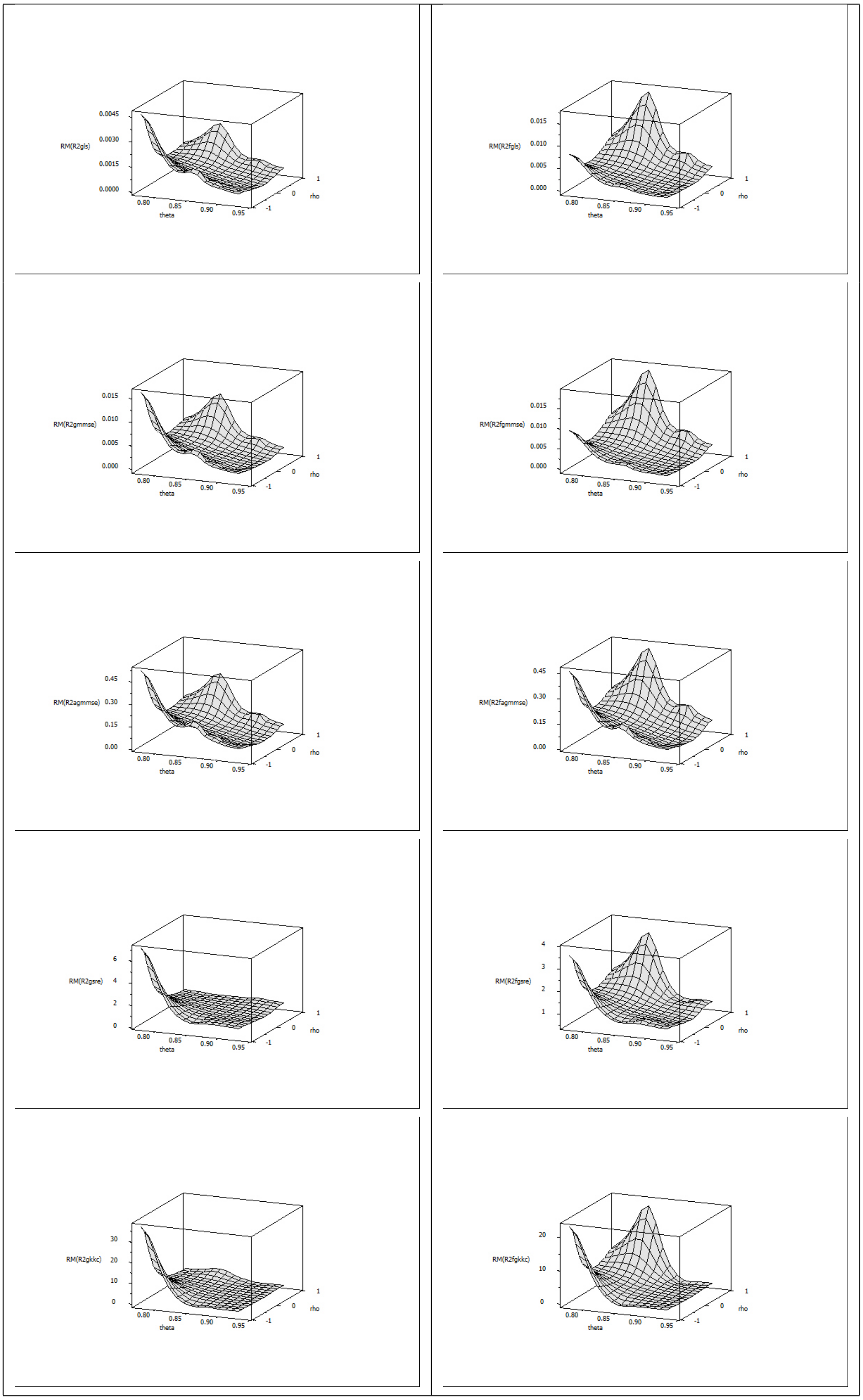

FIGURE 2. Surface plots of relative mean squared error (RM) of R^2_{kk} and R^2_{fkk} when $\sigma^2_v = 10$, $n = 100$ and $p = 8$

6. Conclusion

Considering the setup of multiple linear regression model with random errors not necessarily having an identity covariance matrix, the regression estimates are obtained using DKKE and FDKKE. The goodness-of-fit statistics for DKKE and FDKKE estimators are proposed when the covariance matrix of random errors is known as well as unknown. These statistics can be used in several situations like heteroscedasticity, autocorrelation etc. The first and second order moments up to first order of approximation of such proposed statistics are derived. The empirical findings about the absolute relative bias and relative mean squares errors of these statistics are based on a Monte-Carlo simulation under a first order autoregressive model. The empirical findings from various statistics based on different estimators suggest that the performance depends on the total number of observations of explanatory variables. As this number increases, the performance of statistics improve in the sense of smaller absolute bias and relative mean squared errors. The goodness-of-fit statistics based on GMMSE and GLSE have the smallest absolute bias and the smallest relative mean squared errors, respectively. The same is true for their feasible versions viz., FGMMSE and FGLSE. If the structure of covariance matrix is changed and is not based on a first order autoregressive model, then such ordering may change. The main reporting in this paper is a very general form of a goodness-of-fit statistics. Another interesting finding is that the superiority of an estimator in terms of variance (or mean squared error) may not necessarily carry further over the goodness-of-fit. The proposed statistics capture the model fitting very well even when ϑ is low.

Acknowledgment

Authors are grateful to an anonymous referee whose comments improved the paper.

References

1. 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.
2. A. T. K. Wan, A. Chaturvedi, *Double k-class estimators in regression models with non-spherical disturbances*, Journal of Multivariate Analysis, **79** (2001), no. 2, 226–250.
3. A. Ullah, S. Ullah, *Double k-class estimators of coefficients in linear regression*, Econometrica, **46** (1978), no. 3, 705–722.
4. A. Ullah, S. Ullah, *Errata: Double k-class estimators of coefficients in linear regression*, Econometrica, **49** (1978), no. 2, 554.
5. A. Ullah, V. K. Srivastava, *Moments of the ratio of quadratic forms in non-normal variables with econometric examples*, Journal of Econometrics, **52** (1994), 129–142.
6. A. Chaturvedi, Shalabh, *Risk and Pitman closeness properties of feasible generalized double k-class estimators in linear regression models with non-spherical disturbances under balanced loss function*, Journal of Multivariate Analysis, **90** (2004), no. 2, 229–256.
7. C.-L. Cheng, Shalabh, G. Garg, *Coefficient of determination for multiple measurement error models*, Journal of Multivariate Analysis, **126** (2014), 137–152.
8. C.-L. Cheng, Shalabh, G. Garg, *Goodness of Fit in Restricted Measurement Error Models*, Journal of Multivariate Analysis, **145** (2016), 101–116.
9. J. S. Crämer, *Mean and variance of R^2 in small and moderate samples*, Journal of Econometrics, **35** (1987), 253–266.
10. K. Ohtani, *Minimum mean squared error estimation of each individual coefficients in a linear regression model*, Journal of Statistical Planning and Inference, **62** (1997), 301–316.
11. K. Ohtani, *Exact small sample properties of an operational variant of the minimum mean squared error estimator*, Communications in Statistics (Theory and Methods), **25** (1996), 1223–1231.
12. K. Ohtani, D. E. A. Giles, *The absolute error risks of regression "goodness of fit" measures*, Journal of Quantitative Economics, **12** (1996), 17–26.
13. K. Ohtani, H. Hasegawa, *On small scale properties of R^2 in a linear regression model with multivariate t errors and proxy variables*, Econometric Theory, **9** (1993), 504–515.

14. M. D. Smith, *Comparing approximations to the expectation of quadratic forms in normal variables*, Econometric Reviews, **15** (1996), 81–95.
15. R. A. L. Carter, V. K. Srivastava, A. Chaturvedi, *Selecting a double k-class estimator for regression coefficients*, Statistics & Probability Letters, **18** (1993), no. 5, 363–371
16. R. W. Farebrother, *The minimum mean square error linear estimator and ridge regression*, Technometrics, **17** (1975), 127–128.
17. Shalabh, G. Garg, C. Heumann, *Performance of double k-class estimators for coefficients in linear regression models with non-spherical disturbances under asymmetric losses*, Journal of Multivariate Analysis, **112** (2012), 35–47.
18. T. J. Rothenberg, *Hypothesis Testing in Linear Models when the Error Covariance Matrix is Nonscalar*, Econometrica, **52** (1984), no. 4, 827–842.
19. T. W. Anderson, *An Introduction to Multivariate Analysis*, John Wiley, New Jersey. 2003.

Institute of Statistical Science, Academia Sinica, Taipei, Taiwan, R.O.C.
E-mail address: clcheng@stat.sinica.edu.tw

Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur - 208 016, INDIA
E-mail address: shalab@iitk.ac.in

Department of Statistics, Allahabad University, Allahabad - 211 002, INDIA
E-mail address: anoopchaturv@gmail.com

Received 20.12.2018

СТАТИСТИКА УЗГОДЖЕНОСТІ ДЛЯ УЗАГАЛЬНЕНОГО СТИСКАЛЬНОГО ОЦІНЮВАННЯ

Ч.-Л. ЧЕНГ, ШАЛАБ, А. ЧАТУРВЕДІ

Анотація. Розроблено статистику узгодженості для моделей лінійної регресії, підігнаних за допомогою оцінок стискального типу. Сім'я подвійних оцінок k-класу розглядається як стискальна оцінка, що містить декілька оцінок як окремі випадки. Розглянуто випадки відомої та невідомої коваріаційної матриці похибок, яка за припущенням є відмінною від одиничної матриці. Для сім'ї подвійних оцінок k-класу запропоновано статистику узгодженості, що ґрунтується на ідеї коефіцієнта детермінації у моделі множинної лінійної регресії. Виведено її моменти першого та другого порядку з точністю до першого порядку апроксимації та вивчено її властивості для скінченної вибірки за допомогою методу Монте-Карло.

СТАТИСТИКА СОГЛАСИЯ ДЛЯ ОБОБЩЕННОГО СЖИМАЮЩЕГО ОЦЕНИВАНИЯ

Ч.-Л. ЧЭНГ, ШАЛАБ, А. ЧАТУРВЕДИ

Аннотация. Разработана статистика согласия для моделей линейной регрессии, подогнанных при помощи оценок сжимающего типа. Семейство двойных оценок k-класса рассматривается как сжимающая оценка, которая охватывает несколько оценок как частные случаи. Рассмотрены случаи известной и неизвестной ковариационной матрицы ошибок, которая предполагается отличающейся от единичной матрицы. Для семейства двойных оценок k-класса предложена статистика согласия, основанная на идее коэффициента детерминации в модели множественной линейной регрессии. Выведены ее моменты первого и второго порядка с точностью до первого порядка аппроксимации и изучены её свойства для конечной выборки с помощью метода Монте-Карло.