ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 123, 1992

Jacek Osiewalski[*]

CENTERED AND NONCENTERED VARIANCE INFLATION FACTORS FOR THE OLS ESTIMATOR OF A LINEAR FUNCTION AND FOR THE OLS PREDICTION ERROR[**]

1. INTRODUCTION

Let R denote the correlation matrix for regressors in the classical linear regression model. The diagonal elements r^{ii} of R^{-1} are called "variance inflation factors" (VIF´s), since they indicate how many times larger the variances of the OLS estimators of regression coefficients β_i are for given regressors than in the reference case of $R = I$ (see e.g. Judge et al., 1980, p. 461-462; Mansfield and Helms, 1982, Belsley et al., 1980, p. 93).

In this paper we generalize the concept of VIF to the case of OLS estimation of any given linear function of regression coefficients and to the case of OLS prediction. We consider separately VIF´s based on the usual correlation matrix (for centered regressors in regression with an intercept) and noncentered VIF´s (NVIF´s) based on the noncentered correlation coefficients. Both types of measures give precise numbers indicating an increase (or decrease) of variance of the OLS estimator of a linear function $\gamma = c'\beta$ for given c or of the OLS prediction er-

[*] Lecturer at the Academy of Economics, Kraków.

[**] Written at Tilburg University during a research stay supported by the Netherlands Organization for Scientific Research (NWO). Comments by Mark F. J. Steel on an earlier version are gratefully acknowledged.

ror $f = \hat{y}_* - y_* = x_*'b - (x_*'\beta + u_*)$ for given x_*, but each of the two measures relates to the different reference point (zero correlation coefficients or zero noncentered correlation coefficients).

2. VIF'S BASED ON THE USUAL CORRELATION MATRIX

We consider the linear regression model

$$y = X\beta + u, \quad E(u) = 0, \quad E(uu') = \sigma^2 I_n,$$

where $X = [eZ]$ is $n \times k$ nonrandom of rank k $(k > 2)$ and with a vector of ones (e) as its first column (that is, β_1 is an intercept). Let

$$\bar{z} = [\bar{x}_2 \ \ldots \bar{x}_k]' = \frac{1}{n} Z'e,$$

$$\tilde{Z} = Z - e\bar{z}' = (I_n - \frac{1}{n} ee')Z,$$

$$S = \mathrm{Diag}(s_2, \ldots, s_k), \quad s_i = [\sum_{t=1}^{n} (x_{ti} - \bar{x}_i)^2]^{0.5}$$

$$R = S^{-1}\tilde{Z}'\tilde{Z}S^{-1};$$

that is $\bar{z}$ is a vector of arithmetic means of $k - 1$ nonstochastic regressors (columns of Z), $\tilde{Z}$ is a matrix of deviations from means and R is a correlation matrix (in a purely descriptive sense, because Z is nonrandom).

Since $(X'X)^{-1}$ can be presented in the following form:

$$(X'X)^{-1} = \begin{bmatrix} n & e'Z \\ Z'e & Z'Z \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{n} + \bar{z}'(\tilde{Z}'\tilde{Z})^{-1}\bar{z} & -\bar{z}'(\tilde{Z}'\tilde{Z})^{-1} \\ -(\tilde{Z}'\tilde{Z})^{-1}\bar{z} & (\tilde{Z}'\tilde{Z})^{-1} \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{1}{n} + \bar{z}'S^{-1}R^{-1}S^{-1}\bar{z} & -\bar{z}'S^{-1}R^{-1}S^{-1} \\ -S^{-1}R^{-1}S^{-1}\bar{z} & S^{-1}R^{-1}S^{-1} \end{bmatrix} \tag{1}$$

we can express variances of OLS estimators and predictors in terms of $\bar{z}$, S, R. Precisely, if $g = c'b = c'(X'X)^{-1}X'y$ is the OLS estimator of $\gamma = c'\beta$ $(c \neq 0)$ and $\hat{y}_* = x_*'b$ is the OLS predictor of $y_* = x_*'\beta + u_+$, where $E(u_*) = 0$, $E(u_*^2) = \sigma^2$, $E(u_*u) = 0$, then the partition of c and x_* conformably with $X = [eZ]$:

$$c = [c_1 \quad c_z']', \qquad x_* = [1 \quad z_*']'$$

enables us to write the variance of g and the variance of the prediction error $f = \hat{y}_* - y_*$ in the following forms

$$V(g) = \sigma^2 c'(X'X)^{-1}c = \sigma^2[\tfrac{1}{n}c_1^2 + (c_z - c_1\bar{z})'S^{-1}R^{-1}S^{-1}(c_z - c_1\bar{z})],$$

$$V(f) = \sigma^2[1 + x_*'(X'X)^{-1}x_*] = \sigma^2[1 + \tfrac{1}{n} + (z_* - \bar{z})'S^{-1}R^{-1}S^{-1}(z_* - \bar{z})].$$

Now, if we take as a point of reference a hypothetical set of uncorrelated explanatory variables with the same values of $\bar{x}_i$, s_i $(i = 2, \ldots, k)$, we can define the following variance inflation factors:

$$\mathrm{VIF}(g) = \frac{\tfrac{1}{n}c_1^2 + (c_z - c_1\bar{z})'S^{-1}R^{-1}S^{-1}(c_z - c_1\bar{z})}{\tfrac{1}{n}c_1^2 + (c_z - c_1\bar{z})'S^{-2}(c_z - c_1\bar{z})},$$

$$\mathrm{VIF}(f) = \frac{1 + \tfrac{1}{n} + (z_* - \bar{z})'S^{-1}R^{-1}S^{-1}(z_* - \bar{z})}{1 + \tfrac{1}{n} + (z_* - \bar{z})'S^{-2}(z_* - \bar{z})};$$

they measure how many times larger the variance will be for the given regressors than for uncorrelated ones. In the case of estimating the i-th regression (slope) coefficient β_i $(i = 2, \ldots, k)$ we have $g = b_i$ and VIF(g) reduces to the i-th diagonal element of R^{-1}:

$$\mathrm{VIF}(b_i) = r^{ii},$$

that is, to the variance inflation factor in its form appearing in the literature (see e.g. Judge et al., 1980, p. 461-462, Mansfield and Helms, 1982, Belsley et al., 1980, p. 93). It is well known that $r^{ii} \geqslant 1$ $(i = 2, \ldots, k)$ and the lower bound $(r^{ii} = 1)$ is achieved when the i-th regressor is uncorrelated with the others (see Farrar and Glauber, 1967); that means that correlation between regressors always leads to an increase of variances of the OLS estimators of individual regression (slope) coefficients. Let us stress here that in the general case of the OLS estimation of a linear function of β or in the case of the prediction error, a decrease of variance is also possible and that VIF(·) gives a precise measure of the decrease or increase of variance which is caused by the presence of intercorrelations between regressors. Indeed, VIF(·) can be presented as a ratio

$$\mathrm{VIF}(\cdot) = \frac{a_o^2 + a'R^{-1}a}{a_o^2 + a'a} = \frac{a_o^2 + a'Q\Lambda^{-1}Q'a}{a_o^2 + a'QQ'a}$$

where $\Lambda = \mathrm{Diag}(\lambda_2, \ldots, \lambda_k)$ is a matrix of eigenvalues of R and Q is an orthogonal matrix of eigenvectors of R, so

$$\mathrm{VIF}(\cdot) \overset{>}{\underset{<}{=}} 1 \iff a'Q(\Lambda^{-1} - I_{k-1})Q'a \overset{>}{\underset{<}{=}} 0.$$

Since $\lambda_2, \ldots, \lambda_k$ are positive numbers summing up to $\mathrm{tr}(R) = k - 1$, then for $R \neq I_{k-1}$ some of these eigenvalues must be greater than 1 and some must be less than 1 and the quadratic form $a'Q(\Lambda^{-1} - I_{k-1})Q'a$ is not positive or negative semi definite. This means that VIF(·) can take values greater, equal, or less than 1, which depends on a, that is on $S^{-1}(c_z - c_1\bar{z})$ in the case of estimation or on $S^{-1}(z_* - \bar{z})$ in the case of prediction. The range of values which can be taken by VIF(·) - for a given R and different a_o^2, a - is easy to establish, since for every a:

$$\lambda_{\max}^{-1}a'a \leqslant a'Q\Lambda^{-1}Q'a \leqslant \lambda_{\min}^{-1}a'a,$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are the maximum and minimum eigenvalues of R, respectively. So we have

$$\mathrm{VIF}(\cdot) \geqslant \frac{a_o^2 + \lambda_{\max}^{-1}a'a}{a_o^2 + a'a} \geqslant \frac{\lambda_{\max}^{-1}a_o^2 + \lambda_{\max}^{-1}a'a}{a_o^2 + a'a} = \lambda_{\max}^{-1} > \frac{1}{k-1}$$

(since $1 \leqslant \lambda_{\max} < \mathrm{tr}(R) = k - 1$) and

$$\mathrm{VIF}(\cdot) \leqslant \frac{a_o^2 + \lambda_{\min}^{-1}a'a}{a_o^2 + a'a} \leqslant \frac{\lambda_{\min}^{-1}a_o^2 + \lambda_{\min}^{-1}a'a}{a_o^2 + a'a} = \lambda_{\min}^{-1} < +\infty$$

(since $0 < \lambda_{\min} \leqslant 1$).

3. VIF´S BASED ON NONCENTERED CORRELATION COEFFICIENTS

We consider again the linear regression model

$$y = X\beta + u, \quad E(u) = 0, \quad E(uu') = \sigma^2 I_n,$$

where X is n x k nonrandom of rank k (k ⩾ 2), but not necessarily with a column of ones (the model may or may not have an intercept). Let W denote a diagonal matrix containing the lengths of the columns of X on its diagonal:

$$W = \mathrm{Diag}\left(\sqrt{\sum_{t=1}^{n} x_{t1}^2}, \ldots, \sqrt{\sum_{t=1}^{n} x_{tk}^2}\right);$$

then XW^{-1} is a matrix of standardized, but not centered, values of regressors (the length of each column is 1) and

$$R_N = (XW^{-1})'(XW^{-1})$$

is a k x k matrix of noncentered correlation coefficients between regressors. Let us consider again the OLS estimator $g = c'b$ of $\gamma = c'\beta$ $(c \neq 0)$ and the prediction error $f = \hat{y}_* - y_*$ of the OLS predictor $\hat{y}_* = x_*'b$. We can write their variances as:

$$V(g) = \sigma^2 c'(X'X)^{-1}c = \sigma^2 c'W^{-1}R_N^{-1}W^{-1}c,$$

$$V(f) = \sigma^2[1 + x_*'(X'X)^{-1}x_*] = \sigma^2(1 + x_*'W^{-1}R_N^{-1}W^{-1}x_*).$$

If we take as a point of reference a hypothetical set of orthogonal regressors with the same lengths, we can define the following variance inflation factors (which we will can "noncentered" and denote NVIF):

$$\mathrm{NVIF}(g) = \frac{c'(X'X)^{-1}c}{c'W^{-2}c} = \frac{c'W^{-1}R_N^{-1}W^{-1}c}{c'W^{-2}c},$$

$$\mathrm{NVIF}(f) = \frac{1 + x_*'(X'X)^{-1}x_*}{1 + x_*'W^{-2}x_*} = \frac{1 + x_*'W^{-1}R_N^{-1}W^{-1}x_*}{1 + x_*'W^{-2}x_*};$$

they measure how many times larger the variance will be for given regressors than for orthogonal ones (with the same lengths). In the case of estimating β_i (i = 1, ..., k) we have $g = b_i$ and NVIF(g) reduces to the i-th diagonal element of R_N^{-1}:

$$\mathrm{NVIF}(b_i) = r_N^{ii}.$$

Along the same lines of reasoning as in the previous section, it can be shown that generally NVIF(g) and NVIF(f) can take values greater than, equal to, or less than 1. The range of possible values of NVIF(·) - for a given matrix R_N and different vectors c or x_* - is determined by the eigenvalues of R_N. If $d_1 \geqslant \ldots \geqslant d_k$ denote the eigenvalues of R_N, then

$$0 < d_{min} = d_k \leqslant 1, \quad 1 \leqslant d_{max} = d_1 < k, \quad \sum_{i=1}^{k} d_i = \mathrm{tr}(R_N) = k,$$

and we have

$$\frac{1}{k} < d_{max}^{-1} \leqslant \mathrm{NVIF}(\cdot) \leqslant d_{min}^{-1} < +\infty.$$

Let us note that NVIF(·) is defined for a larger class of linear regression models than VIF(·), since the latter applies only for models with an intercept.

4. A COMPARISON BETWEEN VIF(·) AND NVIF(·)

In order to make such a comparison possible, we must restrict our considerations to the linear model with an intercept. In the case of VIF(·), the hypothetical reference X matrix consists of a column of n ones (e) and mutually uncorrelated regressors with means $\bar{x}_2, \ldots, \bar{x}_k$ and variances $n^{-1}s_2^2, \ldots, n^{-1}s_k^2$, which imply the same squared lengths of columns as for the actual regressors, namely: $s_2^2 + n\bar{x}_2^2, \ldots, s_k^2 + n\bar{x}_k^2$. In the case of NVIF(·), the hypothetical reference X matrix consists of k orthogonal regressors, whose lengths are the same as those of the columns of the actual X matrix. Since there are infinitely many such reference matrices and in (hypothetical) construction one column is chosen arbitrarily (only its length is fixed), we can restrict ourselves to reference matrices with the first column e. Then the remaining k - 1 columns of the reference X matrix are uncorrelated and with zero arithmetic means[1]; fixed lengths and zero means imply that these k - 1 columns have maximum possible variances. When the k - 1 columns of our actual design matrix (except for the first column, e) have zero means, then the reference patterns for VIF(·) and NVIF(·) coincide, since fixing lengths is (under assumption $\bar{z} = 0$) equivalent to fixing variances of regressors; of course, VIF(·) and NVIF(·) coincide in this case.

In order to compare the role of both types of variance inflation factors in the case of collinearity, let us remember that there are two kinds of (linear) near dependencies between columns of $X = [e\ Z]$:

1) dependencies involving e and only one column of Z, that is small variation of a given regressor (see Silvey, 1969, Belsley et al., 1980, p. 90, 170);

2) dependencies involving at least two columns of Z (they make R "almost singular").

By construction, VIF(·) can measure this increase (decrease) of variance which is caused by dependencies of the second type only. On the contrary, NVIF(·) measures an increase (decrease) of variance caused by both types of dependencies. Thus NVIF(·), based on noncentered data, can be a tool for exploring some particular consequences of collinearity. The role of VIF(·), a measure based on centered data, is much more restricted; in the case of collinearity with prevailing dependencies of the first kind (small variation), VIF(·) is misleading as a measure of the consequences of collinearity[2]. In order to avoid misinterpretations of VIF´s and NVIF´s, we should stress that these simple measures add nothing to the explanation of the general statistical consequences of collinearity, as presented by Silvey (1969) (see also Judge et al., 1980, p. 455-458), nor do they substitute the full procedure of detecting collinearity, as

[1] These conditions, that is: $\bar{z} = 0$ and $R = I_{k-1}$, are necessary and sufficient to make X´X diagonal (see (1)).

[2] The problem of centering the data in the context of collinearity is considered in detail by Belsley (1986); he writes about using R (p. 118): "the data correlation matrix [...] will typically produce misleading diagnostic information".

presented by Belsley et al. (1980), Chap. 3. The reason for introducing VIF's and NVIF's is the need for precise numbers indicating the influence of departures from certain "reference patterns" (ideal designs) on the estimation of a particular parameter of interest $\gamma = c'\beta$ or on a particular prediction with fixed x_*. Of course, NVIF(·) can be especially useful in the case of collinearity, but rather in indicating some specific consequences of existing dependencies than in detection of their existence and shape.

In order to compare the values of NVIF(·) and VIF(·) directly, let us write the rations of these measures in the following forms:

$$\frac{\mathrm{NVIF}(g)}{\mathrm{VIF}(g)} = \frac{\frac{1}{n}c_1^2 + \sum_{i=2}^{k} [s_i^{-2}(c_i - c_1\bar{x}_i)^2]}{\frac{1}{n}c_1^2 + \sum_{i=2}^{k} [c_i^2(s_i^2 + n\bar{x}_i^2)^{-1}]},$$

$$\frac{\mathrm{NVIF}(f)}{\mathrm{VIF}(f)} \quad \frac{1 + \frac{1}{n} + \sum_{i=2}^{k} [s_i^{-2}(x_{*i} - \bar{x}_i)^2]}{1 + \frac{1}{n} + \sum_{i=2}^{k} [x_{*i}^2(s_i^2 + n\bar{x}_i^2)^{-1}]}.$$

As it was noticed earlier, if $\bar{x}_2 = \ldots = \bar{x}_k = 0$, then NVIF(·) = = VIF(·). Now let us assume that $\bar{x}_i \neq 0$ for at least one i (i = 2, ..., k).

1. If $c_1 = 0$ (that is, when a linear function under consideration does not involve an intercept β_1), then we obtain

$$\frac{\mathrm{NVIF}(g)}{\mathrm{VIF}(g)} = \frac{\sum_{i=2}^{k} c_i^2 s_i^{-2}}{\sum_{i=2}^{k} c_i^2(s_i^2 + nx_i^2)^{-1}} \geqslant 1;$$

the equality holds only when $c_i = 0$ for all i such that $\bar{x}_i \neq 0$.

2. If $c_1 \neq 0$ and $c_2 = \ldots = c_k = 0$, that is when we are interested in the intercept alone, we have

$$\frac{\mathrm{NVIF}(c_1b_1)}{\mathrm{VIF}(c_1b_1)} = 1 + n \sum_{i=2}^{k} s_i^{-2}\bar{x}_i^2 > 1.$$

3. If for all $i = 2, \ldots, k$ we have $c_i = c_1\bar{x}_i$ in the case of estimation or $x_{*i} = \bar{x}_i$ in the case of prediction, then

$\mathrm{NVIF}(\cdot) < \mathrm{VIF}(\cdot) = 1.$

In other cases the comparison of NVIF(·) and VIF(·) is not as straightforward as above and - generally - NVIF(·) can be greater than, equal to, or less than VIF(·); see values of these measures for $b_1 + b_2$, $x_*'b$ and f in an example found in the next section.

5. AN EXAMPLE

Let us illustrate the generalized definitions of variance inflation factors by the regression equation taken from Theil (1971), Chap. 3, which refers to the consumption of textiles in the Netherlands (1923-1939):

$$\hat{y}_t = \underset{(0.306)}{1.374} + \underset{(0.156)}{1.143}\, x_{t2} - \underset{(0.036)}{0.829}\, x_{t3},$$

where y_t, x_{t2}, x_{t3} denote decimal logarithms of the volume of textile consumption per capita, real income per capita, and the relative price of textiles, respectively; the estimated equation shows the OLS estimates with standard errors in parentheses. In this example:

$$n = 17, \quad \bar{x}_2 = 2.012, \quad s_2 = 0.089, \quad \bar{x}_3 = 1.873, \quad s_3 = 0.385,$$

$$r_{23} = 0.222, \quad r_{N12} = 0.99994, \quad r_{N13} = 0.99876, \quad r_{N23} = 0.99882,$$

where r_{23} is the usual (centered) correlation coefficient and r_{Nij} are the noncentered correlation coefficients. Let us focus here not only on the OLS estimators of individual parameters β_1, β_2, β_3, but also on the OLS estimators of $\beta_2 + \beta_3$, $\beta_1 + \beta_2\bar{x}_2 + \beta_3\bar{x}_3$, $\beta_1 + \beta_2$, and on the OLS predictor corresponding to

$$x_*' = [1 \quad 2.02119 \quad 1.81291],$$

used by Theil (1971), p. 135. Applying definitions of VIF(·) and NVIF(·) we obtain:

$\mathrm{VIF}(b_1) = 0.956,$ $\mathrm{NVIF}(b_1) = 8685,$

$\mathrm{VIF}(b_2) = 1.052,$ $\mathrm{NVIF}(b_2) = 9135,$

$\mathrm{VIF}(b_3) = 1.052,$ $\mathrm{NVIF}(b_3) = 425,$

$$\mathrm{VIF}(b_2 + b_3) = 0.950, \qquad \mathrm{NVIF}(b_2 + b_3) = 4036,$$
$$\mathrm{VIF}(b_1 + \bar{x}_2 b_2 + \bar{x}_3 b_3) = 1, \quad \mathrm{NVIF}(b_1 + b_2\bar{x}_2 + \bar{x}_3 b_3) = 0.33362,$$
$$\mathrm{VIF}(b_1 + b_2) = 0.883, \qquad \mathrm{NVIF}(b_1 + b_2) = 1842,$$
$$\mathrm{VIF}(x_*'b) = 1.093, \qquad \mathrm{NVIF}(x_*'b) = 0.589,$$
$$\mathrm{VIF}(f) = 1.008, \qquad \mathrm{NVIF}(f) = 0.939.$$

In this example, correlation between regressors (measured by r_{23}) is very small and its influence on variances is also small or even negligible. The range of possible values of VIF(·) is narrow:

$$\lambda_{\max}^{-1} = \frac{1}{1.221} = 0.819 \leqslant \mathrm{VIF}(\cdot) \leqslant \lambda_{\min}^{-1} = \frac{1}{0.7786} = 1.284.$$

But in spite of lack of correlation, there is a substantial departure from orthogonality, "caused" by small variation of x_{t2}[3]. This lack of orthogonality gives such large values of NVIF(·) for b_2, b_1, $b_2 + b_3$, and $b_1 + b_2$, but on the other hand it has some positive influence on the variances of $b_1 + \bar{x}_2 b_2 + \bar{x}_3 b_3$, $x_*'b$ and f. Since the eigenvalues of R_N are as follows:

$$d_1 = 2.99835, \qquad d_2 = 0.001592, \qquad d_3 = 0.0000568,$$

the range of possible values of NVIF(·) is very wide:

$$d_{\max}^{-1} = 0.33352 \leqslant \mathrm{NVIF}(\cdot) \leqslant d_{\min}^{-1} = 17612.$$

This example illustrates again the known fact that the consequences of nonorthogonality for the estimation of various parameters and for various predictions can be completely different. The advantage of (generalized) variance inflation factors defined here is that they associate a number with any particular case, and therefore they allow to make quantitative (and not only qualitative) statements about the influences of correlation or nonorthogonality on particular estimators and predictors.

[3] The procedure of detecting collinearity, proposed by Belsley et al.(1980), indicates here a strong dependency which involves only x_{t2} and $x_{t1} \equiv 1$. Condition indexes of XW^{-1} (η_i) and variance-decomposition proportions are as follows:

$$\begin{array}{l c} & \begin{array}{ccc} V(b_1) & V(b_2) & V(b_3) \end{array} \\ \begin{array}{l} \eta_1 = 1 \\ \eta_2 = 43.4 \\ \eta_3 = 230 \end{array} & \left[\begin{array}{ccc} 0.0000 & 0.0000 & 0.0003 \\ 0.0130 & 0.0106 & 0.9851 \\ 0.9870 & 0.9894 & 0.0146 \end{array}\right] \end{array}$$

REFERENCES

Belsley D. A. (1986), *Centering, the Constant, First-Differencing, and Assesing Conditioning*, Chap. 5, [in:] *Model Reliability*, ed. D. A. Belsley, E. Kuh, The MIT Press, Cambridge Mass.

Belsley D. A., E. Kuh, R. E. Welsh (1980), *Regression Diagnostics*, Wiley, New York.

Farrar D. E., R. R. Glauber (1967), *Multicollinearity in Regression Analysis: the Problem Revisited*, "Review of Economics and Statistics", No. 49, p. 92-107.

Judge G. G., W. Griffiths, R. C. Hill, T. C. Lee (1980), *The Theory and Practice of Econometrics*, Wiley, New York.

Mansfield E. R., B. P. Helms (1982), *Detecting Multicollinearity*, "The American Statistician", No. 36, p. 158-160.

Silvey S. D. (1969), *Multicollinearity and Imprecise Estimation*, "Journal of the Royal Statistical Society", B 31, p. 539-552.

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

Jacek Osiewalski

WSPÓŁCZYNNIKI ZWIĘKSZENIA WARIANCJI DLA ESTYMATORA MNK FUNKCJI LINIOWEJ I DLA BŁĘDU PREDYKCJI

Niech R oznacza macierz współczynników korelacji między zmiennymi objaśniającymi klasycznego modelu regresji liniowej

$y = X\beta + u$.

Elementy przekątniowe macierzy R^{-1} nazywane są "współczynnikami zwiększenia wariancji" (ang. variance inflation factors, VIF's), ponieważ informują ile razy większe są wariancje estymatorów MNK parametrów regresji β_i, przy danej macierzy X, niż w idealnym przypadku R = I.

W artykule uogólniamy pojęcie współczynnika zwiększenia wariancji (VIF) na przypadek estymacji MNK dowolnej ustalonej funkcji liniowej parametrów modelu oraz na przypadek predykcji za pomocą predyktora MNK. Rozważamy osobno współczynniki zwiększenia wariancji oparte na zwykłej macierzy korelacyjnej (tj. na scentrowanych wartościach zmiennych objaśniających w przypadku regresji z wyrazem wolnym) i niescentrowane współczynniki zwiększenia wariancji, oparte na niescentrowanych współczynnikach korelacji.

Oba rodzaje mierników dostarczają dokładnych liczb wskazujących wzrost (lub spadek) wariancji estymatora MNK funkcji liniowej $\gamma = c'\beta$ dla danego c lub błędu predykcji $f = \hat{y}_* - y_* = x_*'b - (x_*'\beta + u_*)$ dla danego x_*, ale każdy z tych dwóch mierników (VIF i NVIF) odwołuje się do innego punktu odniesienia (zerowe współczynniki korelacji lub zerowe niescentrowane współczynniki korelacji).