5. A REVIEW OF STUDENT TEST PROPERTIES IN CONDITION OF MULTIFACTORIAL LINEAR REGRESSION

Florin-Marius PAVELESCU*

Abstract

Having in view previous contributions of the author related to the impact of collinearity on the estimated values of parameters of multifactorial linear regressions, in this paper the correlation between Student and Fisher test is emphasized and a correction of the standard computation of the Student test is proposed, in order to increase the respective test relevance and to detect the occurrence of "statistical illusions" determined by collinearity. Also, the impact of adding a new explanatory variable in the linear regression equation is analyzed and the conditions in which such a step is efficient are determined.

Key-words: Correlation between Fisher test and Student test computed values, Coefficient of alignment to collinearity hazard, Corrected Sudent Test, Interdependence between adjusted coefficient of determination and corrected Sudent Test, Anticollinearity

JEL Classification: C51, C52.

1. Correlation between Fisher and Student Test computed values

In a unifactorial linear regression model, $y=a_1 + b_{1k}*x_k$, (1), respectively, the estimated value of parameter b_{1k} is:

$$b_{1k} = \frac{D(y)}{D(x_k)} * R(x_k; y) \quad (2)$$

where: D(y) = standard deviation of the resultative variable observed values;

$D(x_k)$ = standard deviation of the explanatory variable observed values;

$R(x_k;y)$ = coefficient of correlation between the explanatory variable x_k and the resultative variable.

The computed value of the Student test for parameter b_{1k} - t_{b1k}, respectively, is:

* The Institute of National Economy, Bucharest, e-mail: pavelescu.florin @yahoo.com.

$$t_{b1k} = \frac{b_{1k}}{D(u_{1k})} \quad (3)$$

where: $D(u_{1k})$ = standard deviation of errors in case of a unifactorial linear regression.

$$D(u_{1k}) = \frac{D(y) * (1 - R^2(x;y))^{1/2}}{(m-2) * D(x_k)} \quad (4)$$

where: m = number of observations

So it may be written:

$$t_{b1k} = \frac{(m-2)^{1/2} * R(x;y)}{(1 - R^2(x;y))^{1/2}} \quad (5)$$

Having in view the computation formula for the Fisher test for a unifactorial linear regression (F_1) it results that : $t^2_{b1k} = F_1$ (A. Isaic-Maniu, C. Mitruţ, C. Voineagu, 1996).

For a linear regression with n explanatory variables, expressed as: $y = a_n + \sum_{k=1}^{n} b_{nk}$, (6), the computation formula of the Student test corresponding to explanatory variables (t_{bnk}) is:

$$t_{bnk} = \frac{b_{nk}}{D(u_{nk}) * (d_{kk})^{1/2}} \quad (7)$$

However, the estimated value of parameters b_{nk} can be written: (F.M. Pavelescu, 1997):

$$b_{nk} = b_{1k} * T_{nk} \quad (8)$$

where: T_{nk} = coefficient of alignment to collinearity hazard related to explanatory variable x_k.

We notice that the estimated value of parameter b_{nk} may be seen as the product of the estimated value in case of a unifactorial linear regression between the resultative variable and the respective explanatory variable and the coefficient of alignment to collinearity hazard related to the explanatory variable x_k. In these circumstances, we may define the value obtained in case of a unifactorial regression as a **parameter proper estimation value.** Due to the existence of coefficients of alignment to collinearity hazard, the results obtained in multifactorial regressions have to be seen as **parameters derived estimation values** (F.M. Pavelescu, 2004).

The coefficient of alignment to collinearity hazard related to the explanatory variable x_k (T_{nk}) may be written as a ratio of two determinants (F. M. Pavelescu, 1986)[1], respectively:

---

[1] *In the article published in 1986, in the context of Cobb-Douglas production function parameters estimation, I thought that the factor T_{nk} shows the way in which the output dynamics alligns itself to the analysed production factor dynamics. I had taken into account that the values of factor T_{nk} are lower than unit, when they are positive. But, afterwards, when I re-examinated the formula in the general case, and especially in the case of trifactorial regression, I drew the*

$$T_{nk} = \frac{(R_{j1}, R_{j2}...R_{jk-1}, r_{jk}, R_{jk+1}...R_{jn})}{R_{jk}}$$ j=1..n, (9) where: R_{jk} = the coefficient of correlation between the explanatory variables x_j and x_k.

$$r_{xj,xk} = \frac{R(x_j;y)}{R(x_k;y)} \qquad (10)$$

where: $R(x_j;y)$ = coefficient of correlation between the explanatory variable x_j and the resultative variable;
$R(x_k:y)$ = coefficient of correlation between the explanatory variable x_k and the resultative variable.
The ratio r_{jk} can be defined as a coefficient of correlation, mediated by the resultative variable, between the explanatory variables x_j and x_k, related to the explanatory variable x_k. (F.M. Pavelescu, 2005).
Taking into account the computation formula, the factor d_{kk} can be expressed as:

$$d_{kk} = \frac{(R_{jk})_n}{(R_{ji})_{n-1}} \quad (i \neq k) \qquad (11)$$

Also, it may be written:

$$d_{kk} = 1 - c_k \qquad (12)$$

where: c_k = the degree of collinearity induced by adding the explanatory variable x_k in the linear regression equation.

Therefore, $t_{bnk} = [m-(n+1)]^{1/2} * \frac{R(x_k;y)}{(1-R_n^2)^{1/2}} * T_{nk} * (1-c_k)^{1/2}$ (13), equivalent with:

$$t_{bnk} = (nF_{m,n})^{1/2} * \frac{R(x_k;y)}{R_n} * T_{nk} * (1-c_k)^{1/2} \qquad (14)$$

where: $F_{m,n}$ = the computed value of the Fisher test in a regression with m observations and n explanatory variables;

R_n^2 = the coefficient of determination of a linear regression model with m observations and n explanatory variables.

$$F_{m;n} = \frac{m-(n+1)}{n} * \frac{R_n^2}{1-R_n^2} \qquad (15)$$

where: m = number of observations

From the previously presented formula it can be concluded that the computed value of the Student test is sensibly influenced by the computed value of the Fisher test. Under

---

conclusion that the values of factor T_{nk} are mainly influenced by collinearity, measured by the absolute values of coefficients of correlation between the explanatory variables. As the number of explanatory variables grows, the influence of the values of coefficients of correlation between the resultative variable and each of the explanatory variable on the value of factor T_{nk} tends to decrease.

the circumstances of a strict independence between the explanatory variables, all the coefficients of alignment to collinearity hazard are equal to 1, and the collinearity induced by each explanatory variable is null.

Because $R_n^2 = \sum_{k=1}^{n} R^2(x_k; y) * T_{nk}$ (16) (F. M. Pavelescu, 2003) in this case we get the following relationship: $\sum t_{bnk}^2 = nF_{m;n}$ (17). In other words, **the computed value of the Fisher test represent the arithmetical mean of the squares of the computed values of the Student test, if the explanatory variables are strictly independent.**

2. The reasons for correction of the Student test standard computation methodology

The presence of collinearity determines a move off from the relationship between Fisher and Student tests, previously presented. Also, it produces a series of constraints in the use of the Student test.

On the one hand, the respective test cannot be computed if the degree of collinearity is greater than 1. On the other hand, we consider that a linear regression model may not be validated if in the estimation negative coefficients of alignment to collinearity hazard appear. In fact, the existence of negative coefficients of alignment to collinearity hazard reflects the surpassing of a critical level of collinearity, compatible with obtaining relevant results in the parameter estimation. In the context of a calculated value strong polarization of coefficient of alignment to collinearity hazard, taking into account only the absolute values of the Student test, it may lead to the situation where "statistical illusions" are considered as very good estimations for the proposed model.

In order to avoid confusions between truly relevant estimated values and "statistical illusions", we propose a correction of the Student test computation formula, having in view the absolute value of the coefficient of correlation between the resultative variable and the respective explanatory variable observed values.

As a consequence, the computation formula of the Corrected Student test is:

$$CST_{bnk} = (n * F_{m,n})^{1/2} * \frac{/R(x_k; y)/}{R(x_k; y)} * T_{nk} * (1 - c_{nk})^{1/2} \quad (18)$$

equivalent with $$CST_{bnk} = t_{bnk} * \frac{/R(x_k; Y)/}{R(x_k; Y)} \quad (19)$$

where: t_{bnk}= the computed values of the Student test under the circumstances of the standard methodology.

The proposed methodology emphasizes the sign of the coefficient of alignment to the collinearity hazard related to the analyzed explanatory variable. As a consequence, **a first condition for the validation of the parameters estimated values of a**

multifactorial linear regression equation is that the Corrected Student test computed value is positive in the case of all explanatory variables[2].

3. Impact of adding a new explanatory variable to the computed values of Student test

The adding of a new explanatory variable in the linear regression equation determines an increase in the coefficient of determination value, meaning that the resultative variable behavior is better explained. However, due to the decrease in the freedom degree, one may see the increase in the tabled values that define the relevance of each estimated parameter.

As a consequence, it is important to determine the conditions that have to be fulfilled in order to ensure the increase in the Student test value as a result of adding the new explanatory variable.

The ratio of the calculated values of Student test related to explanatory variable x_k in a linear regression equation with (n+1) explanatory variables ($t_{bn+1,k}$) to n explanatory variables ($t_{bn,k}$) in conditions of m observations, respectively, is:

$$\frac{t_{bn+1,k}}{t_{bnk}} = \left(\frac{m-n-2}{m-n-1}\right)^{1/2} * \left(\frac{1-R^2_{n+1}}{1-R^2_n}\right)^{1/2} * \frac{T_{n+1,k}}{T_{n,k}} * \left(\frac{1-c_{n+1}}{1-c_n}\right)^{1/2} \qquad (20)$$

It can be demonstrated that the value of the expression $\frac{(m-n-2)*(1-R^2_n)}{(m-n-1)*(1-R^2_{n+1})}$ is greater than 1, if $\left(\frac{R^2{}_{n+1}-R^2_n}{R^2_n}\right) \geq \frac{1}{m-n-1} * \frac{1-R^2_n}{R^2_n}$ (21). One should note that the fulfillment of the above-mentioned condition leads to an increase in the adjusted coefficient of the determination value (F.M Pavelescu, 2004). In other words, **by adding a new explanatory variable to the linear regression equation, an increase in the adjusted coefficient of correlation value has a favorable influence on the Student test computed value.**

As a rule, $\frac{T_{n+1,k}}{T_{n,k}} * \frac{1-c_{n+1}}{1-c_n} < 1$ (22), due to collinearity increase.

If we denote by $a = \frac{(m-n-2)*(1-R^2_n)}{(m-n-1)*(1-R^2_{n+1})}$ (23) and by $b = \frac{T_{n+1,k}}{T_{n,k}} * \frac{1-c_{n+1}}{1-c_n}$ (24), we

[2] *The condition that all coefficients of alignment to collinearity hazard had to be positive in order to validate the results of the linear regression may be formulated also as: the computed values of the standard Student test ($t_{bn,k}$) and the coefficient of correlation between resultative variable and the analysed explanatory variable ($R(x_k;y)$) must have the same sign.*

can detect three situations in which the ratio $\frac{t_{bn+1,k}}{t_{bn,k}}$ may found itself, if collinearity increases:

1. a*b<1, with a<1, and b<1. The Student test calculated values decrease as a result of a value decrease in the adjusted coefficient of determination and an increase in collinearity.
2. a*b<1, with a>1, and b<1. The Student test calculated values decrease because the favorable influence generated by the increase in the adjusted coefficient of determination is dominated by the unfavorable influence of the increase in collinearity.
3. a*b>1, with a>1, and b<1. The Student test calculated values increase because the favorable influence generated by the increase in the adjusted coefficient of determination dominates the unfavorable influence of the increase in collinearity.

The exception from this rule may occur, for example, when we pass from a unifactorial to a bifactorial linear regression. In this case, $(1-c_{2k})=1$. (25)

In the case of a bifactorial linear regression, with explanatory variables x_1 and x_2, and y as resultative variable, the coeficients of alignment to collinearity hazard are:

$$T_{x1} = \frac{1 - r_{x1,x2} * R(x_1;x_2)}{1 - R^2(x_1;x_2)} \quad (26) \text{ and } T_{x2} = \frac{r_{x1,x2} - R(x_1;x_2)}{r_{x1,x2} * (1 - R(x_1;x_2))}, \text{ respectively.} \quad (27)$$

Usually, the product $r_{x1,x2} * R(x_1;x_2)$ is positive and the regression is validated if $/r_{x1,x2}/ > /R(x_1;x_2)/$, the coefficients of alignment to collinearity hazard being positive and below unit.

But there are cases when the coefficient of correlation between explanatory variables absolute value is close to zero, but its sign is contrary to the sign of the coefficient of correlation between the two explanatory variables, mediated by the resultative variable, $r_{x1,x2} * R(x_1;x_2) < 0$, respectively. As a result, $T_{x1}>1$, and $T_{x2}>1$ and **we can define this situation as a relation of anticollinearity** between the explanatory variables. In this case, the coefficient of determination of the bifactorial regression is higher than the sum of the coefficients of determination of unifactorial regression with either of the two explanatory variables.

As a consequence, if between the two explanatory variable there is a relationship of anticollinearity, three situations can be detected for the ratio $\frac{t_{b2,k}}{t_{b1,k}}$, respectively:

1. a*b>1, with a>1, and b>1. The Student test calculated values increase as a result of value increase of the adjusted coefficient of determination and the anticollinearity.
2. a*b<1, with a<1, and b>1. The Student test calculated values decrease because the unfavorable influence generated by the value decrease of the

adjusted coefficient of determination dominates the favorable influence of anticollinearity.

3. 3) a*b>1, with a<1, and b>1. The Student test calculated values increase because the unfavorable influence generated by the value decrease of the adjusted coefficient of determination is dominated by the favorable influence of anticollinearity.

Therefore, it results that the possibilities for a growth of the Student test computed value as a consequence of adding a new explanatory variable are quite limited. Therefore, the efficiency of such a step is just partial in many cases. Usually, an increase in the coefficient of determination value is obtained, and at a first sight a better explanation of the behavior of the resultative variable of the regression model, but at the same time the relevance of the estimated parameters is diminished.

4. A numerical example. Estimation of GDP elasticities related to manual and non-manual components of the employed population in Romania and Slovakia

We will further illustrate the previous theoretical considerations with a numerical example. For this purpose, parameters were estimated and a series of statistical tests were computed for some production functions that reveal the links between the GDP dynamics and those of "white collar" (non-manual professions) and "blue collar"(manual professions) components of the employed population in Romania and Slovakia, during the period 1995-2002, considering 1994 as a base year.

The production functions whose parameters were estimated are of Cobb-Douglas type and Kmenta type, respectively.

$\ln Y=\ln A_2+\alpha_2*\ln L_w+\beta_2*\ln L_b$ (Cobb-Douglas production function related to white and blue collar components of the employed population)

$\ln Y=\ln A_3+\alpha_3*\ln L_w+\beta_3*\ln L_b+\chi_3*\ln^2(L_w/L_b)$ (Kmenta[3] production function related to white and blue collar components of employed population)

Also, three unifactorial regressions were estimated in order to determine the proper values of elasticities of GDP related to the factors taken into account in Cobb-Douglas and Kmenta type production, respectively:

$$\ln Y=\ln A_w+\alpha_1*\ln L_w$$
$$\ln Y=\ln A_b+\beta_1*\ln L_b$$
$$\ln Y=\ln A_{w/b}+\chi_1*\ln^2(L_w/L_b)$$

where:

$\ln Y$= natural logarithms of GDP indices;

$\ln L_w$= natural logarithms of the "white collar" component of employed population;

[3] *This type of production function was proposed by J. Kmenta in 1967, in order to estimate indirectly the constant elasticity of substitution between labour and capital. In fact, Kmenta's function represents a restricted form of translog (transcendental logarithmic) production function defined in 1971 by Christensen, Jorgensen and Lau.*

lnL_b= natural logarithms of the "blue collar" component of employed population;

A_w, A_b, $A_{w/b}$ A_2 A_3= residual-compensation factors in the case of the above-mentioned regressions;

α_1= proper elasticity of GDP related to the "white collar" component of the employed population;

α_2, α_3 = elasticity of GDP related to the "white collar" component of the employed population in case of Cobb-Douglas and Kmenta type production functions;

β_1= proper elasticity of GDP related to the "blue collar" component of the employed population;

β_2, β_3= elasticity of GDP related to the "blue collar" component of the employed population in case of Cobb-Douglas and Kmenta type production functions;

χ_1= proper elasticity of GDP related to translog of "white collar"/"blue collar" ratio indices;

χ_3= elasticity of GDP related to translog of "white collar"/"blue collar" ratio indices in case of Kmenta type production function.

The data presented in Table 1 show that the estimated proper elasticity of GDP related to the "white collar" component of employed population is negative and very low in absolute value in the case of Romania and positive and relatively high in the case of Slovakia. In obtaining these results, a major contribution was that of the intensity of correlation between the resultative and explanatory variable. In Romania, the GDP dynamics was practically independent from the dynamics of the "white collar" component of employed population, while in Slovakia the above-mentioned correlation may be appreciated as moderate. Therefore, the Corrected Student test computed value is 0.0969 for Romania and 1.1382 for Slovakia.

Table 1

Estimated parameters of functions $lnY=ln A_w+\alpha_1*ln L_w$ and the results of some statistical tests for Romania and Slovakia during the period 1995-2002

Country	Romania	Slovakia
ln A_w	0.0495	-0.0047
α_1	-0.0813	3.4751
Coefficient of determination	0.0016	0.1776
Adjusted coefficient of determination	-0.1648	0.0405
Standard Student Test computed value for α_1	-0.0969	1.1382
Coefficient of correlation between ln Y and ln L_w	-0.0395	0.4214
Corrected Student Test computed value for α_1	0.0969	1.1382

It can be noticed that the estimated proper elasticity of GDP related to the "blue collar" component of the employed population is negative in both countries (Table 2). The absolute value of the respective indicator is lower than 1 in Romania and greater than 1 in Slovakia. Like in the case of the "white collar" component of the employed population, the GDP dynamics is stronger correlated with the explanatory variable in Slovakia than in Romania. Also, the "blue collar" component of the employed population is stronger correlated with GDP dynamics than in the case of the "white collars". Thus, we obtain for Corrected Student test the computed values of 1.1751 for Romania and of 2.4140 for Slovakia.

Table 2

Estimated parameters of functions $\ln Y= \ln A_b+\beta_1*\ln L_b$ and the results of some statistical tests for Romania and Slovakia during the period 1995-2002

Indicator	Romania	Slovakia
$\ln A_b$	0.0346	0.1718
β_1	-0.2688	-1.3293
Coefficient of determination	0.1871	0.4927
Adjusted coefficient of determination	0.0516	0.4082
Standard Student Test computed value for β_1	-1.1751	-2.4140
Coefficient of correlation between ln Y and $\ln L_b$	-0.4325	-0.7019
Corrected Student Test computed value for β_1	1.1751	2.4140

The estimation of the GDP proper elasticity related to translog of the "white collar"/"blue collar" ratio indices leads to positive values in both countries. Also, we should note that absolute values of the above-mentioned pointer are greater than 1 (Table 3). Surprisingly enough, the coefficients of correlation between the resultative variable and the explanatory variable in a translog form take higher values than in the case of the previously estimated functions. As a consequence, the Corrected Student test computed values are 1.5688 for Romania and 3.2269 for Slovakia.

Table 3

Estimated parameters of functions $\ln Y= \ln A_{w/b}+\chi_1*\ln^2 (L_w/L_b)$ and the results of some statistical tests for Romania and Slovakia during the period 1995-2002

Country	Romania	Slovakia
$\ln A_{w/b}$	0.0311	0.1262
χ_1	1.5134	9.0598
Coefficient of determination	0.2961	0.6338
Adjusted coefficient of determination	0.1788	0.5728
Standard Student Test computed value for χ_1	1.5688	3.2269
Coefficient of correlation between ln Y and $\ln^2 (L_w/L_b)$	0.5442	0.7961
Corrected Student Test computed value for χ_1	1.5688	3.2269

The estimation of the parameters of the function $\ln Y=\ln A_2+\alpha_2*\ln L_w+\beta_2*\ln L_b$ and, then, the computation of the above-mentioned statistical tests lead to very different results in the two countries. In the case of Romania, the very low correlation between the logarithms of GDP indices and the logarithms of the "white collar" component of the employed population indices, under circumstances of a moderate correlation between the logarithms of GDP indices and the logarithms of the "blue collar" component of the employed population indices, on the one hand, and between the logarithms of the two explanatory indices, on the other hand, determines the occurrence of a negative value for the coefficient of alignment to collinearity hazard and, implicitly, for the Corrected Student test related to parameter α_2 (Table 4). This fact leads to the invalidation of the respective linear regression.

Table 4

Estimated parameters of functions $\ln Y= \ln A_2+\alpha_2 * \ln L_w +\beta_2 * \ln L_b$ and the results of some statistical tests for Romania and Slovakia during the period 1995-2002

Country	Romania	Slovakia
Ln A_2	0.0613	-0.0412
Coefficient of determination	0.2568	0.6874
Adjusted coefficient of determination	-0.0405	0.5614
Coefficient of correlation between ln L_w and lnL_b	0.5860	-0.0270
α_2	0.6696	3.6338
Standard Student Test computed value for α_2	0.6847	1.7597
Coefficient of alignment for α_2	-8.2360	1.0457
Corrected Student Test computed value for α_2	-0.6847	1.7597
β_2	-0.3873	-1.3518
Standard Student Test computed value for β_2	-1.3103	-2.8507
Coefficient of alignment for β_2	1.4412	1.0185
Corrected Student Test computed value for β_2	1.3103	2.8508

In the case of Slovakia, the estimation premises for the respective bifactorial linear regression are very good. The absolute value of the coefficient of correlation between the two explanatory variables is very close to zero and its sign is contrary to that of the correlation between explanatory variables mediated by resultative variable. In this case, **a situation of anticollinearity between the two explanatory variables occurs**. For this reason, both coefficients of alignment to the collinearity hazard are positive and greater than 1. This fact has favorable consequences for the coefficient of determination, adjusted coefficient of determination and Corrected Student test computed values. It can be observed that the value of the adjusted coefficient of correlation is higher in comparison with both unifactorial linear regressions previously computed. Under such circumstances, the Corrected Student test computed values register a significant growth.

The addition of a new explanatory variable to the regression equation, the translog of the “white collar”/“blue collar” ratio indices, respectively, contributes in the case of Romania to a sensible increase not only in the coefficient of determination, but also in the adjusted coefficient of determination. But, the very high absolute value of the coefficient of correlation between the logarithms of the “blue collar” indices and translog of the “white collar”/“blue collar” ratio indices, determines a strengthening of the degree of collinearity. As a consequence, the coefficient of alignment to the collinearity hazard and the Corrected Student test computed values are negative for parameter β_3. Therefore, this trifactorial linear regression may not be validated (Table 5).

Table 5

Estimated parameters of functions $\ln Y = \ln A_3 + \alpha_3 * \ln L_w + \beta_3 * \ln L_b + \chi_3 * \ln^2 (L_w/L_b)$ and the results of some statistical tests for Romania and Slovakia during the period 1995-2002

Indicator	Romania	Slovakia
$\ln A_3$	0.1745	-0.00001
Coefficient of determination	0.5733	0.6919
Adjusted coefficient of determination	0.2534	0.4609
Coefficient of correlation between $\ln L_w$ and $\ln^2(L_w/L_b)$	-0.5364	0.2362
Coefficient of correlation between $\ln L_b$ and $\ln^2(L_w/L_b)$	-0.9803	-0.9567
α_3	0.3196	2.4005
Standard Student Test computed value for α_3	0.3748	0.4558
Coefficient of alignment for α_3	3.9311	0.6908
Corrected Student Test computed value for α_3	3.7480	0.4558
β_3	1.4643	-0.3109
Standard Student Test computed value for β_3	1.3268	-0.0770
Coefficient of alignment for β_3	-5.4480	0.2339
Corrected Student Test computed value for β_3	-1.3268	0.0770
χ_3	8.1714	6.4903
Standard Student Test computed value for χ_3	1.7227	0.2600
Coefficient of alignment for χ_3	5.3993	0.7164
Corrected Student Test computed value for χ_3	1.7227	0.2600

In the case of Slovakia, adding a third explanatory variable to the linear regression equation prove to be inefficient to a considerable extent. Due to the fact that, like in the case of Romania, a very strong correlation between the logarithms of the “blue collar” indices and translog of the “white collar”/“blue collar” ratio indices can be identified. As a result, a very small growth of the coefficient of determination and a decrease in the adjusted coefficient of determination are registered.

All the three coefficients of alignment to collinearity hazard are positive, so the first condition for the regression validation is fulfilled. But in comparison with the bifactorial linear regression, the Corrected Student test computed values for elasticities of GDP related to the “white collar” and the “blue collar” components of the employed population, respectively, are sensibly diminished.

It is worth mentioning that the coefficients of alignment to the collinearity hazard are rather in contradiction with the contribution to the avoidance of collinearity consequences. Thus, the highest value is attributed to the translog of “white collar”/“blue collar” ratio indices, which represents the explanatory variable that mainly induces the collinearity, due to the strong correlation in absolute value (more than 0.95) with the logarithm of the “blue collar” indices.

The previous numerical examples show that, for multifactorial linear regression equations, the presence of multicollinearity has a noticeable impact on the computed values of statistical tests used in the standard methodology to validate the estimation results. The increase in collinearity creates serious constraints on the effectiveness of an increase in explanatory variable numbers. The above-mentioned phenomenon determines not only a decrease in the Standard Student test computed values, but also creates premises for “statistical illusions”, due to the negative coefficients of alignment to the collinearity hazard. Therefore, it is necessary to use a **corrected version of the Student test**, which is able to detect the sign of the coefficient of alignment to the collinearity hazard. Also, it is important to have in mind that possibilities for an increase in the Corrected Student test computed values as a result of adding a new explanatory variable may occur only if the collinearity induced by the respective variable is very low, or some of the explanatory variables are in the situation of anticollinearity and the computed value of the coefficient of determination in the initial stage is relatively small.

References

C. Allen, S. Hall, (ed.), (1997), Macroeconomic modelling in a changing world, New York: John Willey &Sons.

L.R. Christensen, D.W. Jorgensen and L.J. Lau, (1971), “Conjugate duality and the transcendental logarithmic production function”, Econometrica, 39.

C. Dougherty, (2002), Introduction to econometrics, second edition, Oxford University Press.

W. Greene, (1993), Econometric analysis, McMillan Publishing Company, New York.

A. Isaic-Maniu, C. Mitruț, V. Voineagu, (1996), Statistica pentru managementul afacerilor, Editura Economică, Bucureşti (Statistics for business management, Economica Publishing House, Bucharest.

D. Jula, (2003), Introducere în econometrie, Editura Professional Consulting, Bucureşti (Introduction in econometrics, Professional Consulting Publishing House, Bucharest).

J. Kmenta, (1967), “On the estimation of CES production function”, International Economic Review, 8(2).

G. Maddala, (1988), Introduction to econometrics, McMillan Publishing Company.

F. M. Pavelescu, “Some considerations regarding the Cobb-Douglas production function estimated parameters. A new approach”, Revue Roumaine des Sciences Sociales, Tome 30 No 1-2/1986.

F. M. Pavelescu, (1992), “Considerații privind estimarea parametrilor funcției de producție CES prin metoda Kmenta”*, Revista Română de Statistică* nr. 4-5.

F. M. Pavelescu, (1997), *Progresul tehnologic şi ocuparea forței de muncă,* Editura IRLI, Bucureşti.

F. M. Pavelescu, (2003), "Proprietăți ale coeficientului de determinație" *Revista Română de Statistică* nr. 4.

F. M. Pavelescu, (2004), "Features of ordinary least square (OLS) method. Implications for estimation methodology", *Romanian Journal of Economic Forecasting,* 2.

F. M. Pavelescu, (2005), "Corelația dintre testul Fisher şi testul Student în cazul unei regresii liniare" *Revista Română de Statistică,* 1.

F. M. Pavelescu, (2005), "Impact of collinearity on the estimated parameters and classical statistical tests values of multifactorial linear regressions in conditions of OLS" *Romanian Journal of Economic Forecasting,* 6(2).

E. Pecican, Econometrie, Bucureşti: Editura All

E. Pecican, (1996), *Macroeconometrie. Politici economice guvernamentale şi econometrie.* Bucureşti: Editura Economică.