Analysis of the dynamics of the contribution for social insurance in function by the dynamics of the Gross Domestic Product by econometric modeling

Associate Professor Nicolae Mihăilescu PhD (n.mihailescu@yahoo.com)
„Hyperion" University of Bucharest

Associate Professor Claudia Căpățînă PhD (claudiacapatana@yahoo.com)
„Hyperion" University of Bucharest

Assistant Professor Oana Camelia Iacob PhD *(oanacameliaiacob@yahoo.com)*
Polytechnic University of Bucharest

ABSTRACT

The analysis presented in this article identifies an econometric model of the interdependence of the dynamics of social insurance contribution with the dynamics of Romania's gross domestic product during the period from 2010 to 2017 and the estimated levels of Ministry of Public finances for the years 2018 – 2022.

The developed econometric model has a linear unifactorial form and is supported by adequate statistical support confirming its usefulness for the substantiation of macroeconomic, budgetary and fiscal policy decisions.

Keywords*: contribution to social insurance, gross domestic product, econometric model*

JEL Classification*: E01, E27, E62*

INTRODUCTION

Social contributions are tax obligations through which employers or employees (who in this case can be called taxpayers) are determined to pay monthly percentages of their gross income to public systems, in exchange for insurance.

It should be noted that these contributions are obligatory, by law; they are automatically withheld without the consent of the taxpayer. The contributions are regulated by Law no. 227/2015 regarding the Fiscal Code.

It is mentioned that the law underwent a modification at the end of 2017, given by the Emergency Ordinance no. 79/2017, hereinafter referred to as "GEO 79/2017". Entry into force starting January 1, 2018, GEO no. 79/2017 amended the law so that some of the social contributions that were in the employer's obligation became the employee's obligation.

The contribution for social insurance is dependent on the size of the salary or income for which it is provided by law to calculate a certain quota. Therefore, the economic power of the taxpayers is the measure of the general level of economic development of a country and the gross domestic product is the value expression of this development. A study of the interdependence between the contribution dynamics for social insurance and the dynamics of the gross domestic product has a sustainable motivation and can provide useful information regarding this legality, the speed of changing the value of the social insurance contributions depending on the change of the gross domestic product.

The statistical data based on which the analysis methodology will be performed are presented in Table 1.

The dynamics of the social insurance contribution and the gross domestic product realized between 2010 and 2017 and the levels estimated by the Ministry of Public Finance for the years 2018 - 2022

Table 1

Years	G.D.P.	Contribution for social insurance (consolidated general budget)	
	Amounts millions lei (x)	Amounts millions lei (y)	% of GDP
I. Achievements 2010	533,881.1	45,703.8	8.6
II. Achievements 2011	565,097.2	50,637.6	9.0
III. Achievements 2012	595,367.3	51,658.3	8.7
IV. Achievements 2013	637,456.0	54,383.2	8.5
V. Achievements 2014	668,143.6	57,585.4	8.6
VI. Achievements 2015	712,658.5	57,616.5	8.1
VII. Achievements 2016	762,341.8	61,270.2	8.0
VIII. Achievements 2017	856,700.0	71,705.7	8.4
IX. Estimated 2018 - execution	949,600.0	98,100.8	10.3
X. Proposals 2019 -15 February	1,022,500.0	117,246.1	11.5
XI. Estimates 2020	1,101,000.0	129,757.8	11.8
XII. Estimates 2021	1,178,600.0	143,737.8	12.2
XIII. Estimates 2022	1,261,500.0	158,926.8	12.6

Data source: Ministry of Public Finance

RESEARCH METHODOLOGY

The research methodology of the interdependence of the dynamics of the contribution for social insurance with the dynamics of the gross domestic product is carried out during the following stages:

- The dynamics of the absolute indicators considered as forming an interdependent system are represented graphically - the contribution for social insurance in correlation with the gross domestic product of Romania, from 2010-2022.

- The mathematical form of the model is chosen based on the graphical representation,

- The estimators of the model are defined using the method of the smallest squares and the statistical significance of them is verified using the „*Criterion t*",

- The econometric representation indicators are calculated and the statistical viability of the model is estimated based on a set of statistical tests aimed at: the significance of the correlation ratio with the help of „*Criterion F*", the normality of the distribution of the residual variable with the help of the „*Jarque-Bera Criterion*", the existence of the phenomenon of residual autocorrelation using the „*Durbin-Watson Criterion*" and the residual homoscedasticity phenomenon with the help of the „*White Test*". It also quantifies the „*power*" of the model for calculating predictable levels of social insurance budget revenues depending on the growth of gross domestic product, with the help of „*Theil's coefficient of irregularity / inequality*" as well as by the relative expression of „Estimation of the average error of the regression equation",

- The amount of the social insurance contribution for the year 2023 is estimated, as a point value and as a guaranteed confidence interval with a probability of 95% depending on the value of the gross domestic product.

The methodology used to develop and attest the viability of the contribution dynamics model for social insurance according to the dynamics of the gross domestic product is applied by using the Eviews software.

DEFINING THE MODEL AND CALCULATING THE ECONOMETRIC REPRESENTATION INDICATORS

The analysis of the interdependence of the contribution dynamics for social insurance according to the dynamics of the gross domestic product has as statistical information support the data presented in Table 1.

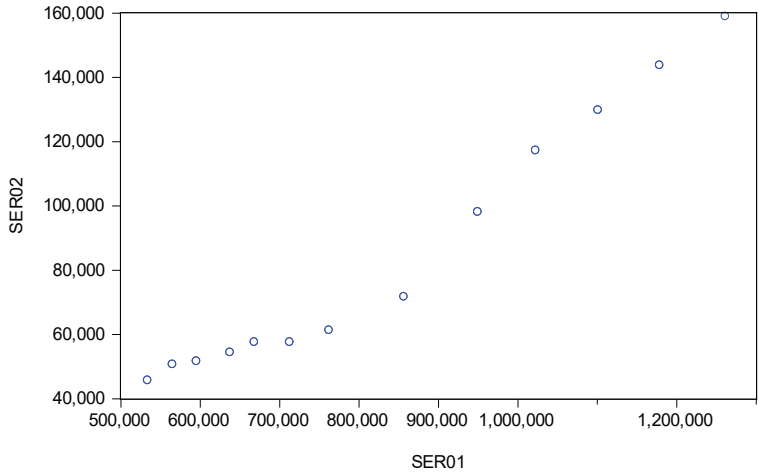
To identify the mathematical form of the econometric model, the graphical representation of the correlation between the variables of the system

under study (Figure. 1) is analyzed. The way in which the „point cloud" is shown in the graph, marked at the intersection of the coordinates of the two variables, provides sufficiently convincing information about the form of their interdependence.

In these conditions, we choose a simple linear regression equation that has the general form of representing the real levels: $y = a + b \cdot x + u$, where y is the endogenous variable (dependent) - the dynamics of the contribution for social insurance, x is the exogenous variable (independent) - the gross domestic product , and u is the residual variable.

Graphical representation of the correlation between the dynamics of the contribution for social insurance according to the dynamics of the gross domestic product

Figure 1



Note: The graph legend (Figure. 1) is explained as follows:
Ser 02 = the contribution dynamics for social insurance (millions of lei)
Ser 01 = dynamics of gross domestic product (millions of lei)

The econometric model developed to meet the proposed knowledge interest is in the form of a simple linear regression equation and the econometric representation indicators are shown in Table 2.

Synoptic table of the system of econometric representation indicators for the linear unifactorial model of the dynamics of social insurance contribution depending on the dynamics of the gross domestic product

Table 2

Dependent Variable: y				
Method: Least Squares				
Sample: 2010 - 2022				
Included observations: 13				
$\hat{y} = a + b \cdot x \rightarrow \hat{y} = -48{,}140.76 + 0.158984 \cdot x$				
Variable	Coefficient	Std. Error	*t*-Statistic	Prob.
x	0.158984	0.009699	16.39110	0.0000
C	-48140.76	8410.500	-5.723888	0.0001
R-squared	0.960668	Mean dependent var: $\bar{y}$		84486.92
Adjusted *R*-squared	0.957092	S.D. dependent var		39939.39
S.E. of regression: $\hat{\sigma}_{y,\hat{y}}$	8273.146	Akaike info criterion		21.02006
$\hat{V}_{y.\hat{y}} = \frac{\hat{\sigma}_{y.\hat{y}}}{\bar{y}} \cdot 100$	9.7922%	Theil Inequality Coefficient		4.1076%
Sum squared resid	(7.53E+08)	Schwarz criterion		21.10697
Log likelihood	-134.6304	Hannan-Quinn criter.		21.00219
F-statistic	268.6680	Durbin-Watson stat		0.425572
Prob (*F*-statistic)	0.000000	Heteroskedasticity Test: White		
Jarque – Bera (*J-B*)	1,492329	*F*-statistic<*F*-table (P = 95%, f_1=2, f_2=10) 0.541624<4.10 – homoskedasticity		q = 0.5979
Prob. (Jarque - Bera)	0,474182	χ^2 -statistic< χ^2 -tabelar (P = 95%, f=2) 1.270587<5.99 – homoskedasticity		q = 0.5298

Table 3 presents the actual levels, the estimated levels and the error term levels or the model residuals that are dimensioned, for each year, by making the difference between the actual levels and the estimated levels of the dependent variable, the value of the social insurance contribution. It is mentioned that the residuals (errors) range is presented in comparison with the size of the average error estimate of the regression equation, $\hat{\sigma}_{y,\hat{y}} = \pm 8{,}273.146$ millions lei , to illustrate their arrangement.

It is also observed that the residuals do not exceed the limits framed ± 2.201 by estimates of the average error of the regression equation, based on the Student's t-distribution, for a significance threshold of 5% arranged bilaterally and 11 degrees of freedom, as confirmed by the graph. from Figure 2.

This statistical finding is able to justify the appreciation that the econometric model of the contribution dynamics for social insurance according to the dynamics of the gross domestic product is a viable analytical representation of the reality.

At the same time, however, there is also a non-alternative positioning of the residuals in relation to the origin, which warns that the errors are affected by the phenomenon of autocorrelation. This graphical finding also has a numerical statistical expression given by the size of the Durbin – Watson coefficient, $DW = 0.425572$.

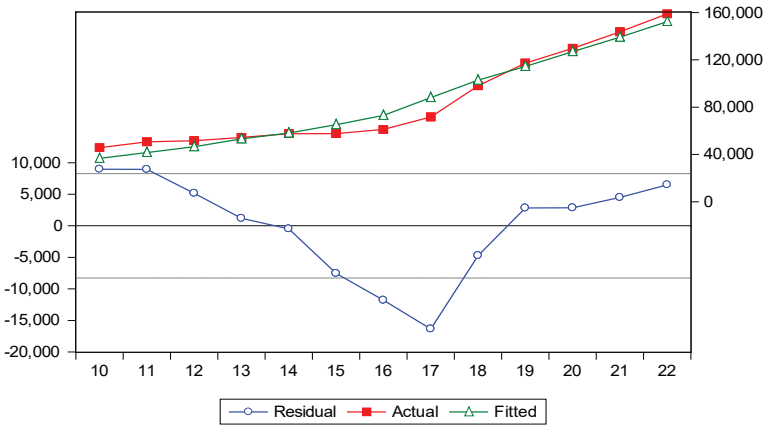
Series of calculation basis levels, of the estimated levels regarding the dependent variable (social insurance contribution) depending on the gross domestic product and the plot of the residual term - (linear unifactorial econometric model)

Table 3

Obs	Actual millions LEI y	Fitted millions LEI $\hat{y}$	Residual millions LEI $u = y - \hat{y}$	Residual Plot $\hat{\sigma}_{y,\hat{y}} = \pm 8{,}273.146$ $-\hat{\sigma}_{y.\hat{y}}$ 0 $+\hat{\sigma}_{y.\hat{y}}$
2010	45703.8	36737.9	8965.86	| . | .* |
2011	50637.6	41700.8	8936.79	| . | .* |
2012	51658.3	46513.3	5145.02	| . | *. |
2013	54383.2	53204.7	1178.48	| . |* . |
2014	57585.4	58083.6	-498.165	| . * . |
2015	57616.5	65160.7	-7544.23	| * | . |
2016	61270.2	73059.6	-11789.4	| * . | . |
2017	71705.7	88061.1	-16355.4	|* . | . |
2018	98100.8	102831.0	-4729.91	| . * | . |
2019	117246.0	114421.0	2825.44	| . |* . |
2020	129758.0	126901.0	2856.88	| . |* . |
2021	143738.0	139238.0	4499.70	| . | * . |
2022	158927.0	152418.0	6508.90	| . | *. |
Sum	1,098,330.0	1,098,330.0	0.00	

Graphic representation of residues, of calculation basis levels and levels estimated for the dynamics of the social insurance' contribution depending on the dynamics of the gross domestic product (Linear unifactorial econometric model)

Figure 2



The graphical form of the estimated levels on the social insurance' contribution, ($\hat{y}$) , is shown in Figure 2 and Figure 3. Thus, the sure trend of growth in a directly proportional report of the social insurance' contribution with the increase of the gross domestic product is identified.

Under the conditions of the unifactorial econometric model that formalizes the correlation of the two variables, it is found that the condition of viability of the model is satisfied in order to calculate forecast levels because „ Theil's Inequality Coefficient" is smaller, compared to the maximum allowed level of 5% (Th = 4.1076%).

It is mentioned that the size "Theil's Inequality Coefficient" can be positioned between zero and one or 100, if expressed as a percentage.
"Theil's Inequality Coefficient " is determined as follows:

$$Th = \frac{\sigma_{y.\hat{y}}}{\sqrt{\frac{\Sigma y^2}{n}} + \sqrt{\frac{\Sigma \hat{y}^2}{n}}} \cdot 100 = 4.1076\,\%$$

$$\text{and } \sigma_{y.\hat{y}} = \sqrt{\frac{\Sigma (y - \hat{y})^2}{n}}$$

It is also considered useful to determine the relative expression of the standard error estimation of the regression equation, which has a statistical significance similar to the „ Theil's Inequality Coefficient „ in order to attest the „power" of the model. This relative estimate is calculated as follows:

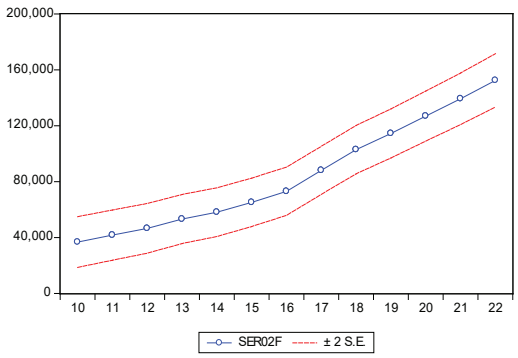
$$\hat{V}_{y.\hat{y}} = \frac{\hat{\sigma}_{y.\hat{y}}}{\bar{y}} \cdot 100 = \frac{8,273.146}{84,486.92} \cdot 100 = 9.7922\%$$

$$\hat{\sigma}_{y.\hat{y}} = \sqrt{\frac{\Sigma(y-\hat{y})^2}{n-k}} = \sqrt{\frac{(7.53E+08)}{13-2}} = \sqrt{\frac{753,000,000}{13-2}} = 8,273.146$$

The relative expression of the standard error estimation of the regression equation provides complementary information that supports the conclusion about the viability of the model (regression equation) for a forecast calculation, and to confirm this fact it is necessary to register a maximum of 10%.

Graphical representation of the dynamic series regarding the estimated levels of the contribution for social insurance according to the gross domestic product and the limits that fit them under the conditions of two estimates of the average error of the regression equation

Figure 3



Forecast: SER02F	
Actual: SER02	
Forecast sample: 2010 2022	
Included observations: 13	
Root Mean Squared Error	7610.188
Mean Absolute Error	6294.934
Mean Abs. Percent Error	9.389326
Theil Inequality Coefficient	0.041076
Bias Proportion	0.000000
Variance Proportion	0.010031
Covariance Proportion	0.989969

INTERPRETING THE RESULTS AND ASSESSING THE VIABILITY OF THE MODEL

The interpretation of the obtained results refers to the significance of the econometric representation indicators on the basis of which the quality and the viability certification of the model are evaluated.

The linear unifactorial econometric model of the contribution dynamics for social insurance according to the dynamics of the gross domestic product has the following analytical form,

$$\hat{y} = -48{,}140.76 + 0.158984 \cdot x$$

and it is confirmed as a model with reserved viability because not all the conditions required by the certification are met:

1- "The Correlation report" has a size close to the maximum limit one (R = 0.980137), in order to obtain the confirmation that there is a very strong correlation of the social insurance contribution with the gross domestic product. Also, by the size of the "Coefficient of determination" (R^2 = 0.960668) it can be specified that 96.0668% of the change in the value of the social insurance contribution is explained by the change of the gross domestic product, the difference up to 100% represents the proportion of the residual component or is motivated by the influence of other factors, not included in the model;

2- the linear unifactorial model of the dynamics of the contribution for social insurance according to the dynamics of the gross domestic product, through the correlation ratio, is viable according to "Criterion *F*", because the obtained result is significantly different from zero, with a probability of over 95% and thus, the existence of a real statistical correlation between the variables of the studied system is validated, considering that *F*-statistic = 268.6680 has a size that exceeds to a large extent the table value of 4.84 (*F*-table = 4.84);

3- the regression coefficient of the model, "*b*", (b = 0.158984) is significantly different from zero (the null hypothesis is rejected), based on "Criterion *t*" with a significance threshold below the maximum limit of 5% for rejection null hypothesis. Under these conditions, the independent (exogenous) variable, the gross domestic product, has a significant influence on the level of the social insurance contribution.

The exposed situation allows us to appreciate that at an increase of one unit (one million of lei) of the independent variable (gross domestic product), the contribution for social insurance increases during the period 2010 - 2022, by 0.158984 units (millions of lei).

Also, the parameter "*a*", the model constant, subject to the same test criterion has the statistical recognition of being significantly different from zero, thus supporting the viability of the regression equation;

4- The "Durbin-Watson statistical coefficient" (DW = 0.425572) has a size that is not positioned in the interval of acceptance of the absence of the autocorrelation phenomenon of the variants of the residual term. The conclusion is formulated based on the Durbin - Watson distribution for a significance threshold of 5%, the number of exogenous variables k'= 1 and the number of observations n = 13;

It is mentioned that the state of self-correlation of the residual terms has the effect of interpreting the following statistical indicators in a reserved way:

- the estimation of the standard deviation of the regression equation has an undervalued and implicit value, the coefficient of determination and the correlation ratio respectively are overvalued. In these conditions, the intensity of the interdependence between the variables of the studied system is of a size marked by a certain distortion;

- "Criterion *t*" used to test the significance of the estimated values of the regression equation parameters is not completely conclusive. In this case the t-statistic values are to a certain extent overvalued, which confirms a better significance of the parameters of the dynamics of the social insurance contribution model depending on the dynamics of the gross domestic product;

5- the relative expression of the "Estimation of the standard error of the regression equation" which has the value of 9.7922% provides the information that the model (the regression equation) has viability, statistically confirmed, for an estimation of forecast estimation, because the size of this indicator does not exceed the acceptance threshold considered 10% restrictive;

6- a statistical significance similar to the one presented by the estimation of the relative standard error of the regression equation is obtained by calculating and interpreting "Theil's Inequality Coefficient" (*Th* = 4.1076%). The econometric model of the dynamics of the social insurance contribution according to the dynamics of the gross domestic product is attested with a corresponding viability from the point of view of this indicator because "Theil's Inequality Coefficient" has a value that does not exceed the allowed limit of 5%;

7- the statistical description of the series of the error term (residual) by the coefficient of asymmetry (Skewness) and the coefficient of flattening (Kurtosis) results "Jarque-Bera statistical coefficient" (*JB* = 1.492329) and the probability related to the JB coefficient (*P* = 47.4182%). This information is based on rejecting the hypothesis of disposition of the values of the error term according to the normal distribution law (*the normality test of the distribution of the residual variable*), because the probability associated with the *JB* coefficient is lower than the critical limit of 60%, which does not support the viability of the model;

8- the test of the existence of the heteroscedasticity state of the errors (residual variable), "White Heteroskedasticity Test", confirms the homoscedasticity property of the linear model of the contribution dynamics for social insurance according to the dynamics of the gross domestic product, based on the two applied statistical criteria, "Criterion *F* "and respectively

"Criterion χ^2 " regarding the auxiliary regression equation of the correlation of the square of the residual levels with the independent variable - the gross domestic product. In these conditions the following assessments can be formulated:

- the error dispersion is constant, does not correlate with the independent variable - the gross domestic product;
- the application of „Criterion t" for verifying the significance of the parameters of the regression equation has statistical support;
- the econometric model gives non-discriminatory importance to all observations related to the residual variable.

The econometric study aimed at elaborating and assessing the viability of the linear unifactorial model of the dynamics of the social insurance contribution according to the dynamics of the gross domestic product of Romania for the period 2010-2022 can be finalized by a synthetic conclusion formulated, with full statistical certainty, that the model has a reserved viability but can be a useful utility to substantiate macroeconomic policy decisions.

The complexity of the international economic and political conjuncture as well as the emergence of unforeseen situations that propagate the change of the economic growth rates can significantly invalidate the punctual forecasts, but there are sustainable statistical determinations that confirm the classification within the space of two guaranteed limits with a certain probability.

The proportion of the contribution for state social insurance from the general consolidated budget in the gross domestic product registered successive increases from 8.6% in 2010, to 10.3% in 2018 and 12.6% in 2020, level estimated by the Ministry of Public Finance (Table 1), by applying some normative measures that facilitated the increase of the degree of collection. It is obvious that in this dynamic situation budgetary funds were provided which gradually allowed to reduce the proportion of subsidies from the state budget for balancing the social insurance budget.

According to Law no. 6/2020 of the state social insurance budget for 2020, the amount of the subsidy from the state budget has the following estimated sizes:

- for the year 2020: 11,309,694 thousand lei,
- for the year 2021: 22,255,813 thousand lei,
- for the year 2022: 22,075,972 thousand lei,
- for the year 2023: 22,779,230 thousand lei,

A calculation of the forecast of the contribution for social insurance in 2023 according to the probable level of the gross domestic product is considered sufficiently safe by extrapolating the model for the period 2010 - 2022, $\hat{y} = -48{,}140.76 + 0.158984 \cdot x$.

If the average annual growth rate of the gross domestic product from 2018 to 2022 is maintained, based on the estimates of the Ministry of Public Finance (*Im* = 1.07), the following scenario regarding the social insurance contribution that will be registered in 2023 can be constructed:

Point value:

$$\hat{y}_{2023} = -48{,}140.76 + 0.158984 \cdot (1{,}261{,}500.0 \cdot 1.07) = 166{,}456.6 \text{ millions lei}$$

and respectively a guaranteed confidence interval with a probability of 95%,
- inferior limit: $li = 166{,}456.6 - 2.201 \cdot 8{,}273.146 = 148{,}247.4$ millions lei
- upper limit: $lu = 166{,}456.6 + 2.201 \cdot 8{,}273.146 = 184{,}665.8$ millions lei

Eroarea limită sau maximă admisă pentru estimarea intervalului de încredere a prognozei contribuţiei pentru asigurări sociale în anul 2023 în funcţie de nivelul probabil al produsului intern brut este rezultatul produsului estimaţiei erorii standard a ecuaţiei de regresie, $\hat{\sigma}_{y;\hat{y}} = \pm 8.273{,}146$ milioane lei, cu factorul de probabilitate (valoarea critică) „*t*" care, în acest caz, este de $\pm 2{,}201$, în condiţiile legii de repartiţie Student (dispunere bilaterală a pragului de semnificaţie $q = 0{,}05$ şi $f = n - k = 13 - 2 = 11$ grade de libertate).

Eroarea limită: $\Delta = \pm t_{q=0{,}05;\ f=n-k=13-2=11} \cdot \hat{\sigma}_{y;\hat{y}} = \pm 18.209{,}19$

The maximum error allowed for estimating the confidence interval of the contribution forecast for social insurance in 2023 depending on the probable level of the gross domestic product is the result of the estimation of the standard error of the regression equation, $\hat{\sigma}_{y;\hat{y}} = \pm 8{,}273.146$ millions lei, with the probability factor (critical value) „*t*" which, in this case, is ± 2.201, under the conditions of the law of Student distribution (bilateral disposition of the significance threshold $q = 0.05$ and $f = n - k = 13 - 2 = 11$ degrees of freedom).

Limit error: $\Delta = \pm t_{q=0.05;\ f=n-k=13-2=11} \cdot \hat{\sigma}_{y;\hat{y}} = \pm 18{,}209.19$ millions lei

CONCLUSIONS

The substantiation of the political decisions regarding the increase of the amount of pensions granted to the persons who according to the law receive these incomes from the state social insurance budget is dependent on the amount of incomes collected in the form of the social insurance contribution of the employees. The economic power of the taxpayers is in turn a result of the economic power of the state expressed synthetically by the statistical indicator, the gross domestic product. The predictable increase of the gross domestic product was also estimated by the amount of the social insurance contribution in 2023. The calculations performed and the resulting econometric model can be a rigorous support to motivate the decision to design the state social insurance budget.

Selective Bibliography

1. Andrei, T.; Bourbonais, R., 2008. *„Econometrie"*, Editura Economică, Bucureşti.

2. Anghel, M.G., 2014, *„Econometric Model Applied in the Analysis of the Correlation between Some of the Macroeconomic Variables"*, Romanian Statistical Review – Supplement/Nr. 1/2014, pp. 88–94.

3. **Anghelache, C., Anghel, M.G., Manole, A**., 2015, *"Modelare economică, financiar-bancară şi informatică"*, Editura Artifex, Bucureşti.

4. **Burghelea, Cristina**, 2014, *„Macroeconomie"*, Editura Transerval, Bucureşti.

5. **Mihăilescu, N.**, 2014 - *„Statistică şi Bazele statistice ale econometriei"*, Editura Transversal, Bucureşti.

6. **Pagliacci, M.; Anghelache G.V.; Pocan I.M.; Marinescu R.T.; Manole A.,** 2011, *"Multiple Regression – Method of Financial Performance Evaluation",* ART ECO – Review of Economic Studies and Research, Editura Artifex, Vol. 2/No.4/2011.

7. **Stancu, S., Andrei, T., Iacob, A.I., Tusa, E.**, 2008 - *„Introducere în econometrie utilizând Eviews"*, Editura Economică, Bucureşti.