Romanian Statistical Review
Revista Română de Statistică

THE JOURNAL OF NATIONAL INSTITUTE OF STATISTICS

3 /2022

www.revistadestatistica.ro

CONTENTS 3/2022

Time Series Analysis by Fuzzy Linear Regression

Richard POSPÍŠIL (richard.pospisil@upol.cz)
Faculty of Arts, Palacký University in Olomouc, Czech Republic

Bogdan OANCEA[1](bogdan.oancea@faa.unibuc.ro)
Faculty of Business and Administration, University of Bucharest,

Miroslav POKORNÝ (miroslav.pokorny@vsb.cz)
Faculty of Electrical Engineering and Computer Science, VSB - Technical University of Ostrava, Czech Republic

Gabriela ŠTEFANOVÁ (gabriela.stefanova@zssenicenh.cz)
Elementary school Senice na Hané, Czech Republic

ABSTRACT

Fuzzy set theory constitutes the theoretical background for abstractly formalizing the vague phenomenon of complex systems. Vague data are defined herein as specialized fuzzy sets, i.e., fuzzy numbers, and a fuzzy linear regression model is described as a fuzzy function with such numbers as vague parameters. We applied a generic algorithm to identify the associated coefficients of the model, and provide both analytically and graphically, a linear approximation of the vague function, together with description of its potential application. We also provide an example of the fuzzy linear regression model being employed in a time series with economic indicators, namely the evolution of the unemployment, agricultural production, and construction between 2009 and 2011 in the Czech Republic. We selected this period since it represents the period when the financial and economic crisis started, and a certain degree of uncertainty existed in the evolution of economic indicators. Results take the form of fuzzy regression models in relation to variables of the time-specific series. For the period 2009-2011, analysis confirmed assumptions held by the authors on the seasonal behaviour of such variables and connections between them. In 2010, the system behaved in a fuzzier manner; hence, relationships between variables were vaguer than otherwise, brought about by factors such as difference in the elasticity of demand, state interventions, globalization, and transnational impacts.

Keywords: *fuzzy set, fuzzy linear regression, genetic algorithms, time series*
JEL Classification: *C22, C51, C49, C65*

1. Coressponding author

1. INTRODUCTION

Regression models are used in engineering practice wherever there is a need to reflect independent variables together with the effects of other unmeasured disturbances and influences. In classical statistical regression, it is assumed that the relationships between dependent variables and independent variables of the model are well-defined and sharp. Although statistical regression has many applications, problems can occur under the following circumstances: the number of observations is inadequate (a small data set); there are difficulties verifying distributional assumptions; vagueness affects connections between input and output variables; ambiguity exists surrounding events or the degree to which they occur; inaccuracy and distortion are introduced by linearization (Shapiro, 2005). In real-world applications, these conditions are non-specific and vague. This is particularly true when modelling complex systems that are difficult to define or measure or where a human element is incorporated into the model.

Fuzzy set theory constitutes the theoretical background for abstract formalization of the vague phenomenon of complex systems. In this paper, we define vague data as specialized fuzzy sets, i.e., fuzzy numbers, and the fuzzy linear regression model is defined as a fuzzy function with such numbers as vague parameters. Estimating the uncertainty of a regression model by applying a fuzzy approach does not require adherence with the presumptions of the classical statistical regression.

Fuzzy regression analysis was performed based on the analysis of the time-specific series (hereinafter referred to as "time series") containing selected macroeconomic variables, which could be of a seasonal character in relation to the national economy. These include indicators of construction production (CPT), agricultural production (APT) and the rate of unemployment (UNT) in the Czech Republic from 2009 to 2011. We selected this specific period since it represents the moment when the financial and economic crisis started, and a certain degree of uncertainty was recorded in the evolution of economic indicators. The data sets used in this study were provided by the Czech Statistical Office (2002).

The choice of macroeconomic variables was based on their seasonal character and interrelationships. While construction and agricultural yield increase during the period of spring to autumn, the level of unemployment in this period generally decreases, although they may evolve in the same period differently; numerous factors affect this, e.g., variation in elasticity of demand for construction and agricultural yield, various levels and forms of state intervention in these segments of the national economy, and the influence of foreign trade and globalization. Along with these forms of the trajectories

of variables pertaining to construction and agricultural production, the unemployment rate may not behave completely normally. The cause of this phenomenon can be seen *inter alia* in the limited elasticity of labour supply, the strong influence of trade unions and the entire labour-related system of social security, which in aggregate distort the labour market. The rest of the paper is organized as follows. Section 2 briefly presents the main concepts used by the fuzzy regression analysis, and section 3 is dedicated to the identification of the fuzzy regression model. Next, we presented the results of the fuzzy regression analysis for the selected time series together with some comments on them in section 4 and 5. The paper ends with a section of conclusions.

2. FUZZY REGRESSION ANALYSIS

The ordinary linear regression model of the investigated system (Seber and Lee, 2003) comprises a linear combination of values of its input variables, as in equation (1) below:

$$Y = A_0x_0 + A_1x_1 + \cdots + A_nx_n = \sum_{i=0}^{n} A_ix_i \qquad i = 0,1,\ldots,n \qquad [1]$$

where $(x_0,\ldots,x_n)$ are input (independent) variables ($x_0 = 1$), $(A_0, A_1,\ldots, A_n)$ are the regression coefficients and Y is the output (dependent) variable.

The conventional regression model assumes that system characteristics are defined as crisp and precise, and deviations between the observed and estimated values of the dependent variable stem from observational errors. The origin of a deviation between the observed and estimated value for the dependent variable may not be significantly caused by poor local variables of the system structure. The causes of these variations do not align with the very sharp nature of system parameters. Such fuzzy phenomenon must also be reflected in the fuzziness of the corresponding parameters of the model.

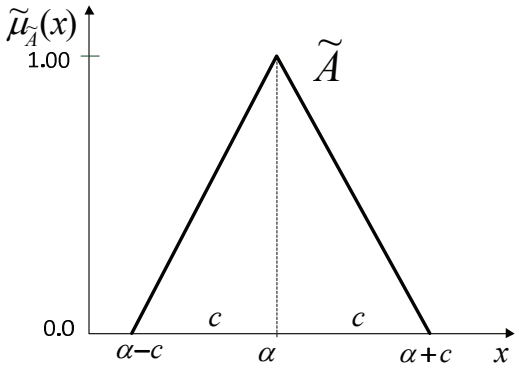
The evolution of the indeterminate regression model occurs through development of the model of vagueness, using the formalization of uncertainty rather than numerical intervals (Ishibushi and Tanaka, 1990; Poleshchuk and Komarov, 2012). Regression models reflecting the vagueness of the modelled systems are called fuzzy regression models (Buckley et al., 2008), (Heshmaty and Kandel, 1985), (Kacprzyk and Fedrizzi, 1992), (Pokorný, 1993), (Shapiro, 2005), (Tansu, 2012). The indeterminate nature of the fuzzy regression model is represented by the fuzzy output values $\tilde{Y}$ and the fuzzy regression coefficients $\tilde{A}$ in the form of specialized fuzzy sets, i.e., fuzzy numbers (Negoita, 2000), (Novák et al., 1999). The form of the fuzzy linear regression model is given by equation (2) below:

$$\tilde{Y} = \tilde{A}_0 x_0 + \tilde{A}_1 x_1 + \ldots + \tilde{A}_n x_n = \sum_{i=0}^{n} \tilde{A}_i x_i \qquad i = 0,1,\ldots,n \quad [2]$$

where $(\tilde{A}_0, \tilde{A}_1, \ldots, \tilde{A}_n)$ are fuzzy regression coefficients (fuzzy numbers); the fuzzy number $\tilde{A}$ is defined by its triangular shape membership function $\mu_{\tilde{A}}(x)$ (see figure 1);

Triangular membership function of fuzzy number $\tilde{A}$

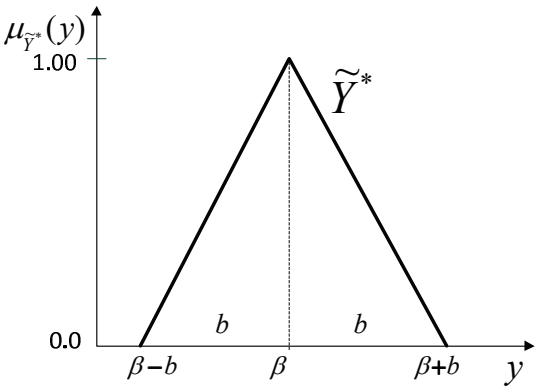
Figure 1



where α is the mean value (core) of fuzzy number $\tilde{A}$ and c is half of the width of the carrier bearing $\tilde{A}\{\alpha, c\}$. The output variable $\tilde{Y}$ of fuzzy regression model 2 is a fuzzy number defined by the triangular membership function (see figure 2). The estimated value $\tilde{Y}^*$ is defined in the form $\tilde{Y}^*\{\beta, b\}$, respectively. The observed value $\tilde{Y}^0$ is denoted in the form $\tilde{Y}^0\{y^0, d\}$ (see Figure 3);

Triangular membership function of fuzzy numbers $\tilde{Y}^*$

Figure 2



where the estimated value β is the mean value (core) of estimated output fuzzy number $\tilde{Y}^*$ and b is half of the width of the carrier bearing $\tilde{Y}^*\{\beta, b\}$.

The parameters β, d are computed applying the principles of fuzzy arithmetic (Mordeson and Nair, 2001); the mean value β is given by equation (3):

$$\beta = \alpha_0 x_0 + \alpha_1 x_1 + \ldots + \alpha_n x_n = \sum_{i=0}^{n} \alpha_i x_i \quad [3]$$

and fuzziness *b* is given by equation (4):

$$b = c_0|x_0| + c_1|x_1| + \ldots + c_n|x_n| = \sum_{i=0}^{n} c_i|x_i| \quad [4]$$

where fuzzification of the examined value y^0 is conducted via fuzzy interval d.

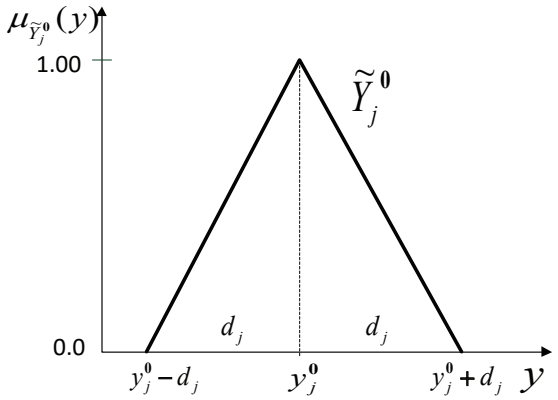
3. IDENTIFICATION OF THE FUZZY REGRESSION MODEL

3.1 Observed output variable y^0 fuzzification

We employed a version where the input variables x comprise crisp numbers, while the observed values $\tilde{Y}^0$ are triangular fuzzy numbers to define the type of the fuzzy regression model. Thus, fuzzy number $\tilde{Y}^*$ is taken as the estimated value and fuzzy number $\tilde{Y}^0$ is the observed value for the model output variable. The fuzziness d_j of observed fuzzy value $\tilde{Y}_j^0$ at step observation *j* is determined via the values gauged at steps (*j+1*) and (*j-1*), respectively (see figure 3).

Triangular membership function of fuzzy numbers $\tilde{Y}^0$

Figure 3



This means that the fuzzy number $\tilde{Y}_j^0$ is of an unequal triangular type; values for d_j are calculated by equation (5):

$$d_j = \frac{1}{2}\left|y_{j+1}^0 - y_{j-1}^0\right| \qquad [5]$$

3.2 Computation of Fuzzy Regression Coefficients $\tilde{A}$

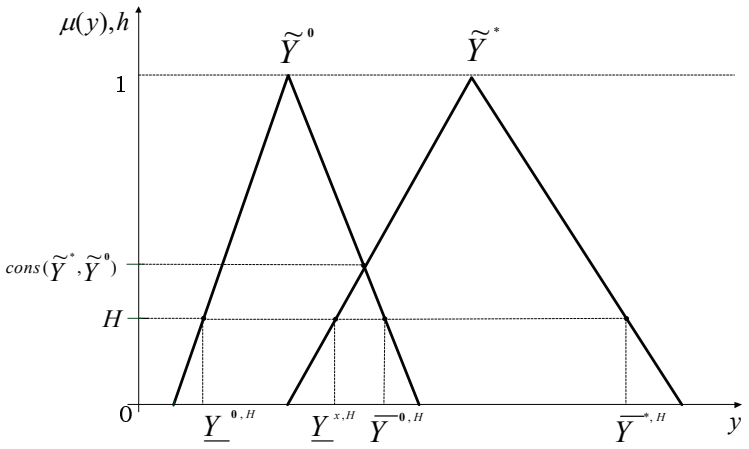
Finding values α and c as search parameters of fuzzy regression coefficients $\tilde{A}_i$ (see Figure 1) is an optimization problem.

The fit of the linear regression fuzzy model to the given data is determined through the Bass-Kwakernaak index *H* (see Figure 4) (Cetintav and Zdemir, 2013), (Kacprzyk and Fedrizzi, 1992). The adequacy of the observed and estimated values is conditioned by equation 6 – the maximum intersection (consistency) of two fuzzy sets; the estimated values for $\tilde{Y}^*$ and the observed ones for $\tilde{Y}^0$ must exceed the set value *H* (see Figure 4).

$$\max_y\{\mu_{\tilde{Y}^0}(y) \wedge \mu_{\tilde{Y}^*}(y)\} = \mathrm{Cons}\left(\tilde{Y}^0, \tilde{Y}^*\right) \geq H \qquad [6]$$

Adequacy of Linear Regression Model

Figure 4



A good estimation of $\tilde{Y}^*$ of the observed output value $\tilde{Y}^0$ is only forthcoming if relation (6) is fulfilled.

Relation (6) is satisfied under the following conditions:

$$\underline{Y}^{,H} \leq \overline{Y}^{0,H} \quad [7]$$

$$\underline{Y}^{0,H} \leq \overline{Y}^{,H} \quad [8]$$

Considering the determined level *H,* the boundary of intervals $Y^{,H}$ and relations (3) and (4), the formulae below are derived:

$$\underline{Y}^{,H} = -(1-H)\sum_{i=0}^{n} c_i|x_i| + \sum_{i=0}^{n} \alpha x_i \quad [9]$$

$$\overline{Y}^{,H} = (1-H)\sum_{i=0}^{n} c_i|x_i| + \sum_{i=0}^{n} \alpha x_i \quad [10]$$

$$\underline{Y}^{,0} = y^0 + (1-H)d \quad [11]$$

$$\overline{Y}^{,0} = -y^0 + (1-H)d \quad [12]$$

Taking $j = 1, 2, \ldots$ and m as the number of observations, conditions (7) and (8) are formulated in final form:

$$\sum_{i=0}^{n}\sum_{j=1}^{m} \alpha_{i,j} x_{i,j} + (1-H)\sum_{i=0}^{n}\sum_{j=1}^{m} c_{i,j} \left|x_{i,j}\right| \geq y_j^0 + (1-H)\overline{d}^0 \quad [13]$$

$$-\sum_{i=0}^{n}\sum_{j=1}^{m} \alpha_{i,j} x_{i,j} + (1-H)\sum_{i=0}^{n}\sum_{j=1}^{m} c_{i,j} \left|x_{i,j}\right| \geq -y_j^0 + (1-H)\underline{d}^0 \quad [14]$$

$$c_{ij} \geq 0 \quad [15]$$

The requirement for adequacy of the estimated and observed values (6) is complemented by the need for minimal total uncertainty of the identified fuzzy regression function:

$$\sum_{i=0}^{n}\sum_{j=1}^{m} c_{i,j} \rightarrow \min, \qquad i = 0, 1, \ldots, n, \quad j = 1, 2, \ldots, m \quad [16]$$

where $i = 1, 2, \ldots$, and n is the number of input values of the regression function, $j = 1, 2, \ldots$ and m is the number of observations.

Then, the optimization problem can be set:
a) minimization of fuzzy model vagueness (16)
b) under condition (6)

To solve the minimization problem (16) under condition (6), numerous authors have employed the linear programming method (Arabpour and Tata, 2008), (Kacprzyk and Fedrizzi, 1992). Nevertheless, a genetic algorithm method is used herein to solve this problem (16). The main reason for this

is that the authors are oriented towards unconventional methods of artificial intelligence to substantiate their quality and efficiency in solving complex tasks (Oancea et al, 2021). Genetic algorithms are a representative of evolutionary methods, and their higher computational complexity is tackled nowadays by the availability of cheap high-performance computers. Thus, they are widely used to solve optimization problems and are commonly employed to identify fuzzy regression models geared towards discerning optimal fuzzy regression coefficients as triangular fuzzy numbers.

Identification of fuzzy regression coefficients - the fuzzy numbers $\tilde{A}_0, \tilde{A}_1, \ldots, \tilde{A}_n$ - was split into two tasks:

a) discerning the mean value (core) α_i of fuzzy number $\tilde{A}_i$ and

b) discerning c_i as a half of the width of the carrier bearing $\tilde{A}_i = \{\alpha_i, c_i\}$.

These tasks were solved via a genetic algorithm. Identification of α_i and c_i was carried out first. Thus, the optimization of the fuzzy linear regression model comprised a two-step process where two genetic algorithms, designated GA1 and GA2, were applied.

For identification of the mean value (core) α_i of fuzzy number $\tilde{A}_i$, minimization of the fit function J_1 is defined in the form below:

$$\min J_1 = \min \frac{1}{m} \sum_{j=1}^{m} \left(y_j^0 - \beta_j\right)^2 \quad [17]$$

where the genetic algorithm GA1 is used. For identification of c_i as a half of the width of the carrier bearing $\tilde{A}_i$, minimization of the fit function J_2 is defined as:

$$\min J_2 = \min \sum_{j=1}^{m} \sum_{i=0}^{n} \left| c_{j,i} \right| \quad [18]$$

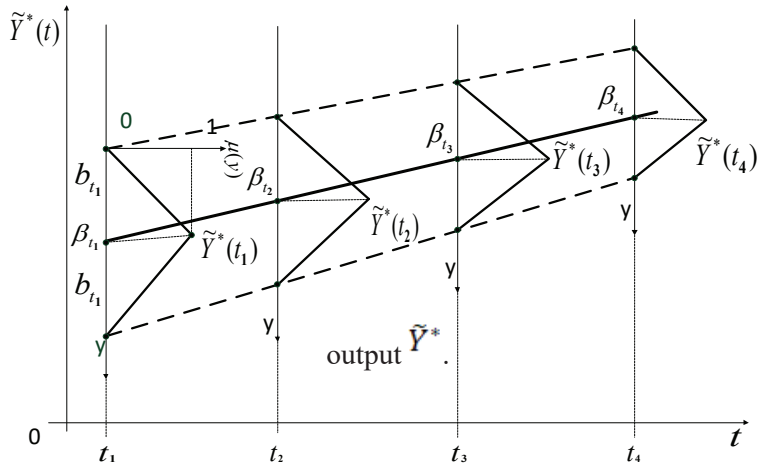
where the genetic algorithm GA2 with three constraints (13), (14), (15) is used. Minimization of the fit function J_2 is based on prior identification of the role of the mean value (core) α_i and utilizes identified values of α_i to determine the width of carrier bearing α_i.

4. ANALYSIS OF TIME SERIES FUZZY REGRESSION

The fuzzy linear regression model can express not only the analytical linear approximation of multivariate functions, but also the size of its uncertainty (vagueness, fuzziness) in the form of an indeterminate potential. Figure 5 shows the graph of a one-dimensional fuzzy regression function together with the appropriate linear approximation and the potential of the estimated fuzzy

One-Dimensional Fuzzy Linear Regression Function

Figure 5



The one-dimensional fuzzy time series regression model can express trends and seasonal cycles. Both are enhanced by the potential that defines the size of the vagueness of the model and defines the range in which the value of the trend and seasonal cycles may fall.

The one-dimensional fuzzy linear regression model of a time series trend is given by the formula

$$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 t \qquad t = 1, 2, \ldots \qquad [19]$$

The value of a seasonal deviation in every month MSD (as a fuzzy number) is calculated for each year $r = 1,2,\ldots, L$ and for each month $k = 1, 2, \ldots, 12$ as the difference between the trend value and the actual value to be estimated:

$$\text{MSD} = \left(\tilde{Y}^0_{r,k} - \tilde{Y}^*_{r,k}\right),\ r = 1,2,\ldots, L,\ k = 1, 2, \ldots, 12 \qquad [20]$$

The central value of fuzzy number MSD is calculated as the difference between the central values $\tilde{Y}^0_{r,k}; \tilde{Y}^*_{r,k}$, and the fuzziness is calculated as the sum of fuzziness of fuzzy numbers $\tilde{Y}^0_{r,k}; \tilde{Y}^*_{r,k}$.

The seasonal cycle is then defined as the time series of 12 seasonal deviations for 12 months. A seasonal deviation for a given month $k = 1, 2, ..., 12$ is calculated as the average value for the month of year $r = 1, 2, ..., L$ of the considered time series.

$$\tilde{Y}_k^* = \frac{1}{L}\sum_{r=1}^{L}\left(\tilde{Y}_{r,k}^0 - \tilde{Y}_{r,k}^*\right);,\ r = 1,2,\ldots, L,\ k = 1, 2, \ldots, 12 \quad [21]$$

For example, the seasonal variation for the first month of January is calculated as the mean of the seasonal variations for January of the considered $r = 3$ years:

$$\tilde{Y}_1^* = \frac{1}{3}\sum_{r=1}^{3}\left(\tilde{Y}_{r,1}^0 - \tilde{Y}_{r,1}^*\right) \quad [22]$$

The values of monthly deviations are calculated as fuzzy numbers. The core of fuzzy number $\tilde{Y}_k^*$ is calculated as the mean difference between the cores, while the uncertainty is calculated as the mean of the sum of fuzziness. Thus, 12 fuzzy numbers are calculated that pass into the timeline of 12 months as a curve of the cores and their potential.

5. ANALYSIS OF SELECTED ECONOMIC VARIABLES

Modelling economic variables with high degree of uncertainty is very difficult, especially during economic and financial crisis. The variability of such variables is subject to several influences, both exogenous and endogenous, some of which are hardly predictable at all or have a prominent degree of fuzzitivity. The relative effect of non-economic influences upon the evolution of the selected economic variables has risen in importance, as various subjects on the market – households and companies – adapt their level of consumption, investment, and savings based on an uncertain future. Apart from rational evaluation of the relevant economic data, they are also under pressure from several influences from the areas of psychology, politics, demographic development, natural circumstances, foreign affairs, etc., and the so-called transactional motive is replaced with the motive of caution.

The time series used in this analysis consist in twelve measured values of the selected variables from 2009 to 2011. The analysed period was selected based on the beginning of the economic crisis, as 2009 was the first year when the crisis fully proceeded throughout the year. The selection of variables was methodically chosen with regard to the mutual interconnectivity and their relative importance in the economy. This was the reason why two primary variables from GDP were analysed (construction and agricultural production) as well as a secondary variable (unemployment), which is in causal relationship to the two previous variables. Both construction and agricultural production are variables with a highly seasonal cycle, which is mirrored in the evolution

of unemployment in both directions following a delay. Simultaneously with this assumption, construction, and agricultural production, however seasonal, may act differently, caused by the obvious differences in the characteristics of these variables. While the elasticity of demand of agricultural production may be very low, it is very high for construction, thus households and companies postpone their consumption and investments for a more favourable period. Therefore, the decrease in the extent of construction works leads to increase in unemployment, yet this is not true for agricultural production. The latter is applicable to the so-called Giffen goods effect, where demand for agricultural production does not rise or fall significantly, as the effect of income fully negates the influence of substitution. Although an increase in unemployment along with a decrease in the level of agricultural production tends to be affected by the seasonal cycle, the increase in unemployment as it relates to the construction sector is usually dictated by decrease in demand for building works.

Construction and agricultural production usually increase during the spring to autumn and undergo decline in the winter. An opposite trend is shown by the unemployment rate, though, which drops in the spring to autumn, and unemployment generally reaches its highest values in the winter (seasonal unemployment). While this is purely conjecture, several certified works and empirical observations support it. Indeed, the economy could have behaved unpredictably in the crisis during the studied years (2009 - 2011). It is interesting to note the positive values for agricultural production when consumers simply cannot significantly reduce demand, and great fluctuations are evident, including negative values in the construction sector. However, for both types of production, despite the diversity, the seasonal behaviour of variables are somewhat apparent, especially if a long-term perspective is taken.

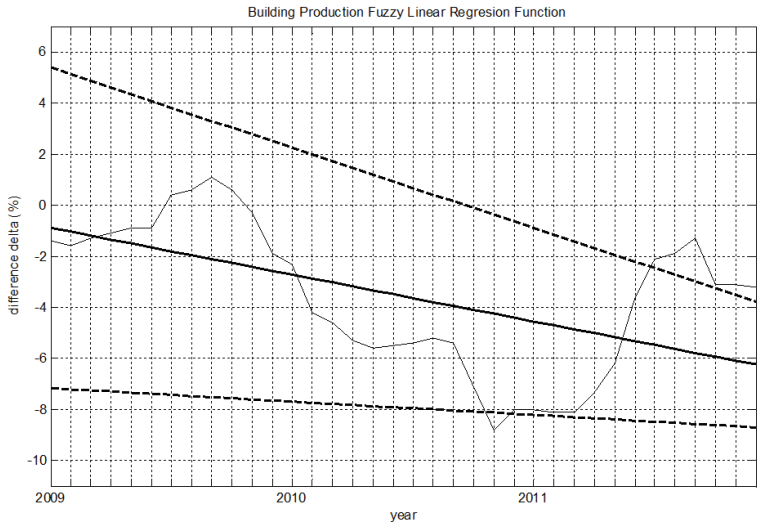
Unemployment behaves in the opposite way, since should production decline (GDP), unemployment goes up and vice versa (due to fluctuation - decrease/increase in demand). Hence, unemployment secondarily shows a seasonal character. It adheres *inter alia* to the so-called Okun law formulated in the 1960s, which states that if there is a decline in GDP by 2%, the unemployment rate grows by 1% or the proportion is approximately 2:1.

Identification of the time series fuzzy regression models was made by applying standard algorithms in Optimtoolbox MATLAB software (Matlab, 2018).

The results are shown as fuzzy regression models of the time series for construction production (CPT; figures 6 and 7), agricultural production (APT; figures 8 and 9) and unemployment (UNT; figures 10 and 11). Figures 6-11 represent their fuzzy trends and fuzzy seasonal cycles. Appropriate regression coefficients of regression functions are presented in the form $A\{\alpha; c\}$

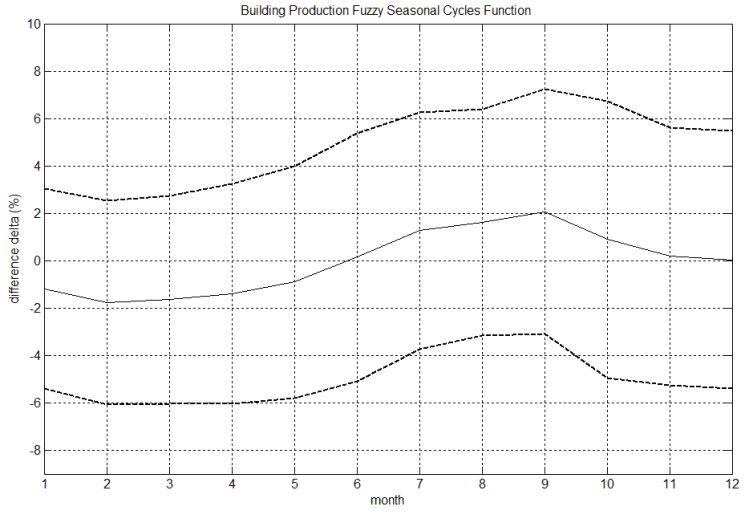
Construction Production - Fuzzy Linear Regression Function

Figure 6



Construction Production - Fuzzy Seasonal Cycles Function

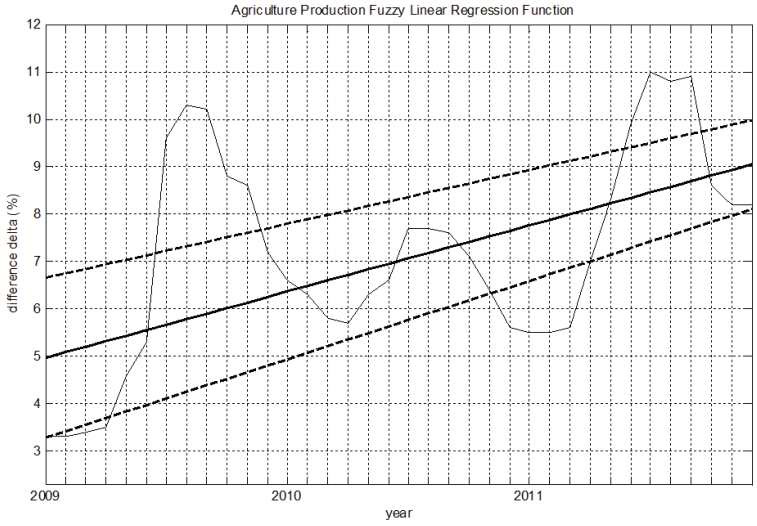
Figure 7



$$A_0\{4.9646; 0.7891\}$$
$$A_1\{0.1157; 0.0243\}$$

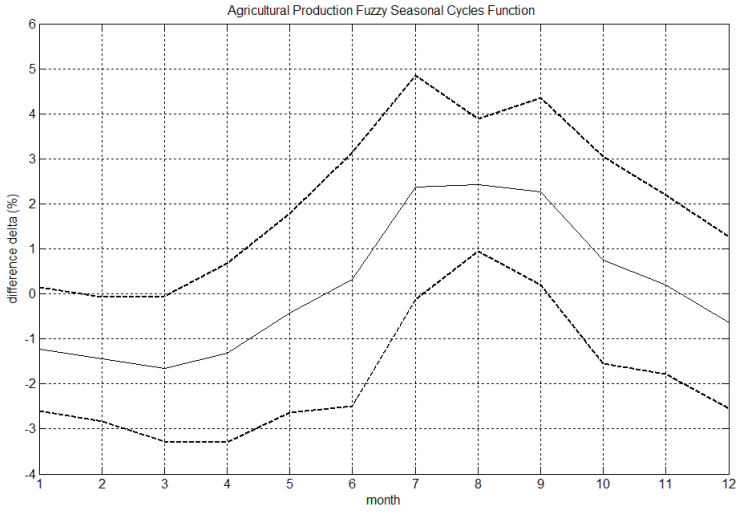
Agricultural Production - Fuzzy Linear Regression Function

Figure 8



Agricultural Production - Fuzzy Seasonal Cycles Function

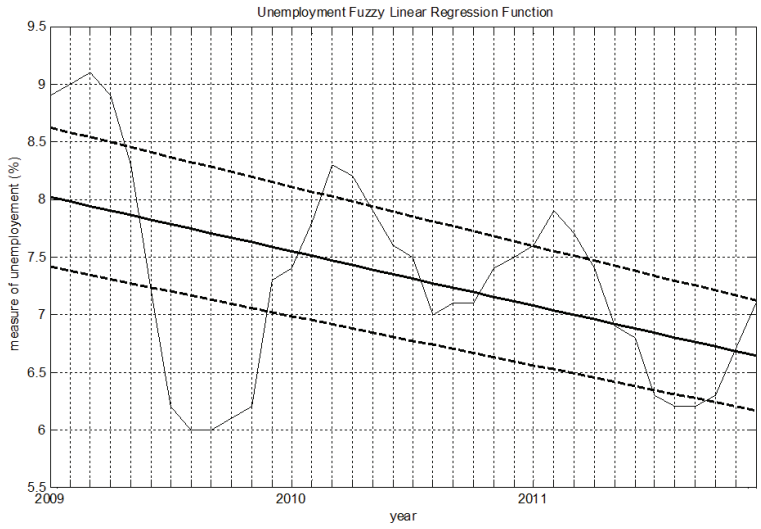
Figure 9



$$A_0\{7.9438; 0.5790\}$$
$$A_1\{-0.0385; 0.0057\}$$

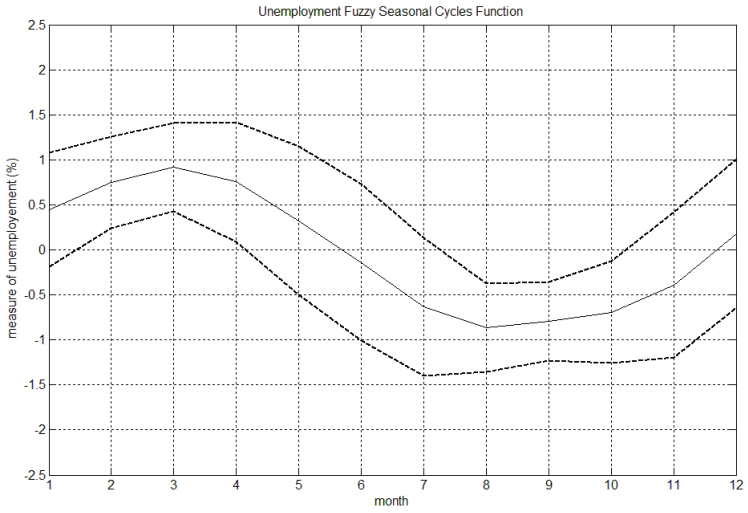
Measure of Unemployment - Fuzzy Linear Regression Function

Figure 10



Measure of Unemployment - Fuzzy Seasonal Cycles Function

Figure 11



The results of the time series analysis of the macroeconomic variables (UNT, CPT and APT) show the interdependence of some of these variables, but in some cases, they also reflect a certain degree of vagueness, i.e., fuzzitivity.

This concerns both the interdependence of CPT and APT variables in relation to UNT, and in some periods the same trend of CPT and APT (2009 and 2011) and their opposite trend in 2010. The dependencies selected above were confirmed by several scientific methods and long-term professional empirical observation (Boeckh, 2010). However, some of the assumptions mentioned in Boeckh (2010) were not confirmed by the presented work for reasons that can be satisfactorily explained.

One of these relates to market failure due to the global economic crisis. This does not refer to a state of stagnation or moderate inflation, but current so-called stagflation, which was once a relatively rare phenomenon. It is a combination of two failures of macroeconomic equilibrium, namely economic stagnation, or rather stagnation of GDP growth, and rising prices (inflation). The existence of this type of failure raises serious national economic problems with an impact on the fiscal and monetary policy of the country and an emphasis on the contradictory nature of these failures, especially the choice of current fiscal expansionary and restrictive monetary instruments of economic policy (Rubin, 2011). What also plays a role here is the global type of economy and thus limited effectiveness of measures at a national level, which is especially true for small and open economies, such as in the Czech Republic.

Another important influence on the variables UNT, CPT and APT over time is referred to as "time lag" in the economy. This is a series of delays resulting from the characteristics of an economic process based on a premise that from the moment when the problem (failure) emerged, while observing the issue with conclusive measurable economic tools (recognition lag), some time (delay) always passes; then there is a particular time interval needed for making a decision and choosing tools for correcting the failure, and a time interval is required to implement the necessary tools, including their positive effect (implementation lag). This fact significantly reduces the efficiency of economic policy and, together with its global character, fundamentally affects economic activity (Friedman, 1994). It also refers to the mutual correlation of all the variables and their existent and proven fuzzitivity.

Another significant circumstance affecting the fuzzitivity of the monitored system is distortion of the market by existing governmental and political interference. In the monitored set of variables, the APT variable is especially affected by agricultural subsidies at national and European levels and deflects the behaviour of particular economic market agents (Mises, 2009). To some extent, this also applies to the UNT variable, which is influenced, for example, by the minimum wage, state employment policy, amount of social benefits and a variety of other interventions that unilaterally deflect the labour market out of the free market. A relatively free market environment exists

only through construction sector. Government interventions that tend to grow definitely increase the vagueness of the variable behaviour of the monitored system.

The intermediate effects of the crisis, which gradually changed within the three observed years from a financial to an economic crisis, also have an indubitable effect on the high fuzzitivity of the investigated system. The crisis concurrently spread slowly from individual market subjects to a crisis of public budgets and state debt crisis. This phenomenon, being much stronger in Eurozone countries than the Czech Republic, exerted an imminent influence upon foreign demand, upon which the Czech economy, being small and open, was very dependent to a certain extent. This was especially the case of agricultural production, semi-products and foodstuffs (i.e., generally APT), and to a smaller extent the export of construction materials, construction workers and investment construction units (generally CPT). Foreign influences, however, tended to follow another route. It was mostly the large import of agricultural production into the Czech Republic, where also typical and traditional agricultural products of both cattle and plant-based production were imported into the Czech Republic, as well as technical and construction materials. This import narrowed the operating space for Czech manufacturers and their supply was limited. In this context, foreign influences have the largest affect upon the variable UNT, as the free movement of the labour force is one of the freedoms of the European free market. The analysed economical areas of APT and CPT are rather less demanding in matters of the labour force qualifications; therefore, they are most affected by the tide of a foreign labour force. This feature cannot be influenced at the national level; thus, it has an imminent influence upon the growth of UNT and is one of the reasons for its high fuzzitivity.

Despite the above-mentioned facts, it is still possible to observe dependencies in the monitored variable set described in this work. In 2009 and 2011, there was a similar trend for CPT and APT seasonal cycles in the summer (from June to September), with a clear decrease in UNT, and in 2011 this trend was even stronger than in 2009. It is a well-known phenomenon of production growth (in this case, CPT and APT) with a parallel decrease of unemployment during summer, or more precisely, with the rise of unemployment during winter, which is known as seasonal unemployment. In 2010, the system behaved fuzzily with an unproven dependency of CPT and APT on UNT. The same year the cycle amplitude of APT was significantly lower than that of CPT, caused by elasticity of demand, i.e. the proportion of change in the demanded quantity and price. Elasticity of APT (agricultural production and foodstuffs) is much lower, sometimes almost zero, compared

to the elasticity of CPT (private and public construction works combined), which has a high elasticity. Therefore, the trend zones and seasonal cycles of APT were significantly narrower than the ones for CPT and sustainably achieved smaller fluctuations. The same was also true for the fuzzitivity of the relation between the APT and UNT variables, which was significantly lower than between CPT and UNT.

Comparing the trend zones for CPT and APT in relation to UNT during the monitored period, the existence of the Okun law can be demonstrated, this being an empirical relationship between the cyclical movements of GDP (herein the CPT and APT variables) and UNT. The law says that if actual GDP drops towards a potential one by 2%, the unemployment rate (UNT) increases by approximately 1%. This relationship applies to the GDP (not only to the sum of CPT and APT); however, but contradictory movement of these variables is also proved by this study. While the APT trend zone sees growth (CPT relatively stagnant), the UNT trend zone decreases, i.e. when production increases, unemployment decreases. This phenomenon can be observed in the variables during the period of 2009-2011, while in the last year of the period the phenomenon shows itself most strongly.

All the variables investigated above have an immediate effect on the fiscal area of the economic policy of the state. While the level of production of actual GDP (parts of which are also APT and CPT) affects the level of tax allocation for public costs, the level of UNT affects the level and rate of their later redistribution. However, the demand for APT and CPT exerts an immediate effect on monetary matters pertaining to the economic policy of the state. Economic entities then react to an anticipated drop in inflation by trying to obtain interest-bearing assets by selling other assets. Through this, they attempt to reduce losses from holding liquid assets that they had obtained by continuous inflation. Such purchases of new assets, however, lead to rise in prices and drop in real pay-off, meaning that even an expected increase in inflation brings about reduction in the interest rate. In economic literature, this effect is called the Mundell-Tobin effect.

From the results of the time series analysis of UNT, CPT and APT herein, the authors demonstrate the interdependence of these variables, and in some points even their high fuzzitivity. This is mainly due to the global nature of the economy, protracted economic crisis, time delays, especially state interventions, and political measures, which influence the free market and national economy.

6. CONCLUSIONS

In classical statistical regression, it is assumed that the relationship between dependent variables and independent variables of a model is well defined and sharp. Although statistical regression has many applications, problems can occur in situations when the number of observations is inadequate (a small data set), or difficulties arise in verifying distributional assumptions, vagueness affects the relationship between input and output variables, ambiguity exists in events or degree to which they occur, or inaccuracy and distortion are introduced by linearization.

In this study we defined vague data as specialized fuzzy sets - fuzzy numbers and devised a fuzzy linear regression model as a fuzzy function with such numbers as vague parameters. Determining the uncertainty of the regression model via a fuzzy approach does not require that the above presumptions are met.

A genetic algorithm was applied to identify the fuzzy coefficients of the model. The linear approximation of the vague function together with its potential was presented analytically and graphically.

Several assumptions, concerning the evolution of the CPT, APT and UNT variables, their seasonality and relationship between them were proved by performing fuzzy regression analysis of the selected variables. In the first (2009) and third (2011) years of observation there was a common and seasonal growth in CPT and APT, while the increase of CPT was lesser in extent due to a higher elasticity of demand for construction, as well as to the full impact of the economic crisis in this segment of economy. The assumption that UNT dropped during the studied period along with the increase of CPT and APT was also confirmed by fuzzy regression analysis.

The fuzzy regression analysis of the time series of CPT, APT and UNT revealed the non-standard behaviour of the monitored variables in 2010. This constituted the third and most profound year of the crisis, and the full influence of the state was evident alongside the huge impact of globalization on the small and open economy of the Czech Republic. Delay certainly played a role here, emerging in the economy during the second studied year (2010). That year the system of indicators behaved fuzzily and the interdependence of CPT and APT on UNT was not proven by the model; moreover, the model behaved in a much vaguer fashion, i.e., fuzzily, in connection with CPT to UNT rather than APT to UNT. The cause of the phenomenon lay in the limited elasticity of demand for agricultural production, or, for example, in the rising price of agricultural commodities throughout the period. State intervention and transnational influences on the APT and UNT variables were so great

that they could be seen as one of the causes of the non-standard and fuzzy behaviour of these variables during the year.

REFERENCES

1. **Arabpour, A.R., Tata, M.**, (2008), *"Estimating the Parameters of a Fuzzy Linear Regression Model"*, Iranian Journal of Fuzzy systems, Vol.5, No. 2, pp.1-19.
2. **Boeckh, J. A.,** (2010), *The Great Reflation, Wiley and Sons*, Inc., USA.
3. **Buckley, J.J., Jowers, L.J.** (2008), *Fuzzy Linear Regression I*. Studies in Fuzziness and Soft Computing, Vol. 22, Springer.
4. **Cetintav, B., Zdemir, F.**, (2013), *LP Methods for Fuzzy Regression and a New Approach* E.Krause, Eds. In Synergies of Soft Computing and Statistics for Intelligent Data Analysis, volume 22, Springer.
5. **Friedman, M.**, (1994), *Money Mischief: Episodes in Monetary History*, First Harvest Edition, USA.
6. **Heshmaty, B., Kandel, A.**, (1985), "*Fuzzy linear regression and its applications to forecasting in uncertain environment*", Fuzzy Sets and Systems no. 15, pp. 159-191.
7. **Ishibuchi, H., Tanaka, H.**, (1990), "*Identification of fuzzy parameters by interval regression models*", Electronics and Communications in Japan (Part III: Fundamental Electronic Science) 73.12, pp. 19-27.
8. **Kacprzyk, J., Fedrizzi, M.**, (Eds.) (1992), *Fuzzy Regression Analysis*. Studies in Fuzziness and Soft Computing, Physica-Verlag HD.
9. **MATLAB,** (2018), 9.7.0.1190202 (R2019b). *Natick, Massachusetts*: The MathWorks Inc.
10. **Mises, L.v.,** (2009), *The Theory of Money and Credit*. New Haven: Yale University Press.
11. **Mordeson, N.J., Nair, P.S.**, (2001), *Fuzzy Mathematics.* Physica-Verlag.
12. **Negoita, C.**, (2000), Fuzzy Sets, New Falcon Publications.
13. **Novák,V., Perfilieva, I., Močkoř, J.**, (1999), *Mathematical Principles of Fuzzy Logic.* Kluwer, Boston.
14. **Oancea B., Pospíšil R., Jula M.N., Imbrișcă C-I**, (2021), "*Experiments with Fuzzy Methods for Forecasting Time Series as Alternatives to Classical Methods*", Mathematics, 9(19):2517. https://doi.org/10.3390/math9192517.
15. **Poleshchuk, O, Komarov, E.**, (2012), *A fuzzy linear regression model for interval type-2 fuzzy sets*, 2012 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS), pp. 1-5, DOI: 10.1109/NAFIPS.2012.6290970.
16. **Pokorný, M.** (1993), *Fuzzy Nonlinear Regression Analysis*. Doctoral Thesis (in Czech) VUT Brno, Faculty of Electrical Engineering, Brno.
17. **Reeves, C.R., Rowe, J.E.,** (2002), *Genetic Algorithms: Principles and Perspectives Kluwer Academic Publishers*, New York.
18. **Rubin, I.S.**, (2011), *The Politics of Public Budgeting*. Chatham: Chatham House, USA.
19. **Seber, G.A.F., Lee, A.J**. (2003), *Fuzzy Regression Analysis*. Wiley Series in Probability and Statistics.
20. **Shapiro, A.F.**, (2005), *Fuzzy regression models*. Available at: http://www.soa.org/research/arch/2006/arch06v40n1-ii.pdf.
21. **Czech Statistical Office**, (2022), https://www.czso.cz/csu/czso/home
22. **Tansu, A.,** (2012), *Fuzzy Linear Regression: Fuzzy Regression*. Lambert Academic Publishing.

Insights on the Impact of Covid-19 Papers Written by Affiliates of Romanian Universities with Medicine Specialization

Lecturer PhD Andreea MIRICĂ (miricaandreea89@gmail.com)
Bucharest University of Economic Studies

Assistant Lecturer, PhD Ionela-Roxana PETCU (roxana.glavan10@yahoo.com)
Bucharest University of Economic Studies

Associate Professor, PhD Traian-Ovidiu CALOTĂ (traian.calota@infofisc.ro)
Titu Maiorescu University, Faculty of Finances, Banks, Accounting and Busineess Administration, Bucharest, Romania

PhD. Candidate Roxana-Violeta PARTAS-CIOLAN (roxana.partas@gmail.com)
Bucharest University of Economic Studies, Bucharest, Romania,

PhD. Candidate Liliana CATRINA[1] (liliana.catrina@gmail.com)
Bucharest University of Economic Studies, Bucharest, Romania,

ABSTRACT

Due to its impact on all human activity, the new virus has therefore aroused increased interest in research. A large number of medical papers related to COVID-19 have been published as a result of the scientific community response to this emerging infectious disease. All measures that combat the spread of the disease came at a cost and the people had to add one more hardship which contributed to the rise in social inequality. Moreover, the costs for internet access, proper equipment to support online education, private healthcare, facemasks, and disinfectants have become a burden for a segment of Romanian society.

In this study, 18 Romanian universities a Medicine bachelor specialisation, both public and private, were analysed to determine the dataset query from Web of Science. Our research aimed to assess the impact of COVID-19 papers written by authors affiliated with Romanian universities with a medicine specialisation, which are

1. Coressponding author

the factors that influence the likelihood of a paper being cited and how the number of citations varies by university type.

The obtained results reveal that University of Medicine and Pharmacy "Carol Davila" from Bucharest records the highest number of published papers, while "APOLLONIA" University of Iași has the highest number of citations in both WOS Core Collection and in all WOS Databases per research. Furthermore, there is a significant difference in the number of citations between papers with and without a PUB MED ID. Logistic models show that neither the type of university nor the property type of the university (public or private) are determinants of the likelihood of an article being cited in the WOS core collection or in all WOS databases. As a consequence, a research paper with PUB MED ID and a higher number of pages, has a higher chance to be cited.

Keywords: *Romania, research assessment measures, bibliometric, COVID-19, research productivity*

JEL Classification: *I200, I230*

1. INTRODUCTION

The COVID-19 pandemic and its consequences led to a broad impact in all human activity. The lack of information and the need to adopt recommendations to tackle the new disease led to an increased activity in the online environment, social media, traditional media and in publicity. Due to the initial adopted measures, as a result of the decrease in mobility and lockdowns (Andrei et al. 2021), people more time with their families at home, but the economy also started to be affected (Albu et al., 2020; Păunescu, and Matyus, 2020) and measures had to be imposed to protect the loss of jobs.

For a while, working from home where possible was favoured, education shifted to online learning and only basic services kept running business as usual. Such changes came suddenly with unforeseen consequences for the economy, challenging the democratic way of life (Raiu and Mina-Raiu, 2022), even questioning the notion of political freedom (Raiu and Juknevičienė, 2021). All measures that combat the spread of the new disease came at a cost (Onofrei et al., 2021) and the people had to add one more hardship which in some of the cases increased the social inequality. In this regard, the costs for facemasks, disinfectants, healthcare, internet access, proper equipment to support online education (Ionescu et al., 2020) became a burden for a population segment in Romania. Not leaving any citizen behind (United Nations, 2015) is the starting point in the analysis of measures that need to be taken so that people are not discriminated, and inequalities are reduced.

The degree of interconnection between people through online social networks such as Facebook, Twitter and many others is unprecedented in human history, and this facilitated the spread of conflicting information (Garrett, 2020) due to existing uncertainties. Limaye et al. (2020) pointed out

that the isolation and physical distancing in the context of reducing the spread of COVID-19 can deepen the use of social media networks as individuals try to maintain connectivity with each other.

This situation has accelerated interest in research and analysis in this field has increased the number of published unreviewed research. Legitimate concerns are also highlighted in the work of Horbach (2020), who studied the length of the publishing process in medical journals before and after the pandemic. The published results reveal that the process is accelerated for publishing coronavirus work, while for the rest, the process remained unchanged.

Although the accelerated publishing process has increased the dissemination of information, the peer review expediting in scientific and medical research adds yet a new pressure to the mechanisms for transmitting information in the current situation. As Bagdasarian et al. (2020) pointed out, many journals provide rapid peer review for COVID-19 manuscripts, encouraging pre-print publication and speeding up the publishing process.

Published papers on COVID-19 can be found in the fields of health, social sciences, economics, education, communication, and new technologies. Recent research of Paun et al. (2020) suggests that the Google Scholar Citations and the Mendeley Citations (reads) are potential predictors of WOS citations.

In this study, we build the database used in our analysis focusing on Romanian universities with bachelor specialisation in Medicine. Therefore, we investigated whether there is an impact of COVID-19 research papers written by authors affiliated with Romanian universities with a medicine specialisation. We examine the number of citations by type of university or indexing in other databases by analysing the factors that influence the likelihood that a paper in this specific database to be cited.

2. METHODOLOGY

For the purpose of this paper, the data was retried from Web of Science on April 30th 2022. The database was queried for author affiliation and topic. In the affiliation field, we entered every Romanian university that has a Medicine bachelor specialisation:

- University of Medicine and Pharmacy "Carol Davila" Bucharest
- "Grigore T. Popa" University of Medicine and Pharmacy
- "Iuliu Haţeganu" University of Medicine and Pharmacy
- Victor Babeș University of Medicine and Pharmacy Timișoara
- Transilvania University of Brașov

- "Lucian Blaga" University of Sibiu
- Oradea University
- "Dunarea de Jos" University of Galați
- University of Medicine and Pharmacy Craiova
- George Emil Palade University of Medicine, Pharmacy, Science, and Technology of Târgu Mureș
- Ovidius University of Constanța
- Ștefan cel Mare University of Suceava
- Titu Maiorescu University
- University of Pitești
- Western University "Vasile Goldiș" from Arad
- "APOLLONIA" University of Iași
- "Constantin Brâncuși" University of Târgu Jiu
- Bioterra University

This list was taken from the open datasets' portal data.gov.ro and contains information about the number of students by bachelor for each university as well as the type of university (public or private). When querying the WOS database, Bioterra didn't return records.

To answer the first research question, a graphical analysis was performed to assess the number of the citations in the WOS core collection as well as in all WOS databases. The characteristics of the article (number of pages, year of publication, article type, whether or not the article has a PUBMED ID, Web of Science Index), as well as universities of the authors (property type, whether the university is multidisciplinary or has exclusively medical specialisations) were considered.

To answer the second research question, the ANOVA technique was applied to test the following set of null hypotheses:

- Null hypotheses 1: There is no significant difference in the number of WOS Core Collection citations between papers with a PUB MED ID and papers without such an ID
- Null hypotheses 2: There is no significant difference in the number of citations in all WOS databases between papers with a PUB MED ID and papers without such an ID
- Null hypotheses 3: There is no significant difference in the number of WOS Core Collection citations between papers belonging to a public university affiliate and a private university affiliate
- Null hypotheses 4: There is no significant difference in the number of citations in all WOS databases between papers belonging to a public university affiliate and a private university affiliate

- Null hypotheses 5: There is no significant difference in the number of WOS Core Collection citations between papers belonging to an affiliate to a university with only medical degree programs and one to a multidisciplinary university
- Null hypotheses 6: There is no significant difference in the number of citations in all WOS databases between papers belonging to an affiliate to a university with only medical degree programs and one to a multidisciplinary university.

The chosen threshold is 5%. In order to answer the third question, two logistic regression models were built:

$$Model\ 1\text{: } \log\left(\frac{P_i}{1-P_i}\right) = a_0 + a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4 + a_5X_5$$

$$Model\ 2\text{: } \log\left(\frac{Q_i}{1-Q_i}\right) = a_0 + a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4 + a_5X_5$$

$P_i - the\ probability\ for\ a\ paper\ to\ be\ cited\ in\ WOS\ Core\ Collection;$
$Q_i - the\ probability\ for\ a\ paper\ to\ be\ cited\ in\ all\ WOS\ Databases;$
$X_1 = 1\ if\ a\ paper\ has\ a\ PUB\ MED\ ID; 0\ otherwise$
$X_2 = Number\ of\ pages$
$X_3 = Year\ of\ publication$
$X_4 = 1\ if\ the\ university\ queried\ is\ multidisciplinary; 0\ otherwise$
$X_5 = 1\ if\ the\ university\ queried\ is\ public; 0\ otherwise$

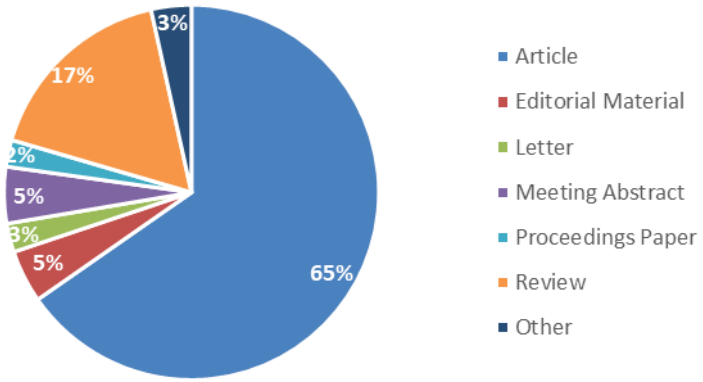
Factors are removed from the model if the z-test probability is higher than 5%.

3. RESULTS

Firstly, the results of the exploratory analysis of the characteristics of the papers taken from the query are presented. 65% of the retrieved papers are research articles followed by reviews (Figure 1). Moreover, over half of the papers are between 1 and 10 pages long, while 36% of them are between 11 to 20 pages long (Figure 2). Most of the papers were written in 2021 (Figure 3) and have a PUB MED ID (Figure 4) meaning that they are present in the PUB MED database. Additionally, approximately half of all the analysed materials were published in journals indexed Science Citation Index Expanded (Table 1).

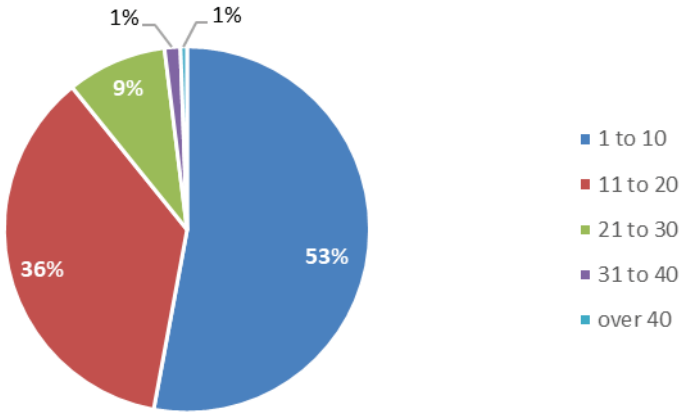
Share of papers by type of document; source: designed by the authors based on WOS data

Figure 1



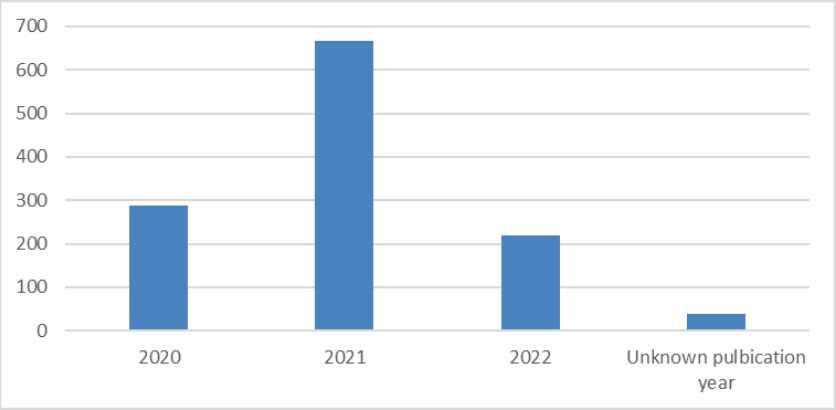
Share of papers by number of pages; source: designed by the authors based on WOS data

Figure 2



Number of papers by year of publication; source: designed by the authors based on WOS data

Figure 3



Number of papers by Web of Science Index; source: designed by the authors based on WOS data

Table 1

Web of Science Index	Number of papers
Conference Proceedings Citation Index - Science (CPCI-S)	22
Conference Proceedings Citation Index - Science (CPCI-S); Conference Proceedings Citation Index - Social Science & Humanities (CPCI-SSH)	3
Conference Proceedings Citation Index - Social Science & Humanities (CPCI-SSH)	3
Emerging Sources Citation Index (ESCI)	231
Science Citation Index Expanded (SCI-EXPANDED)	676
Science Citation Index Expanded (SCI-EXPANDED); Conference Proceedings Citation Index - Science (CPCI-S)	14
Science Citation Index Expanded (SCI-EXPANDED); Social Science Citation Index (SSCI)	215
Social Science Citation Index (SSCI)	46
Social Science Citation Index (SSCI); Arts & Humanities Citation Index (A&HCI)	1

Share of papers by PUB MED status; source: designed by the authors based on WOS data

Figure 4

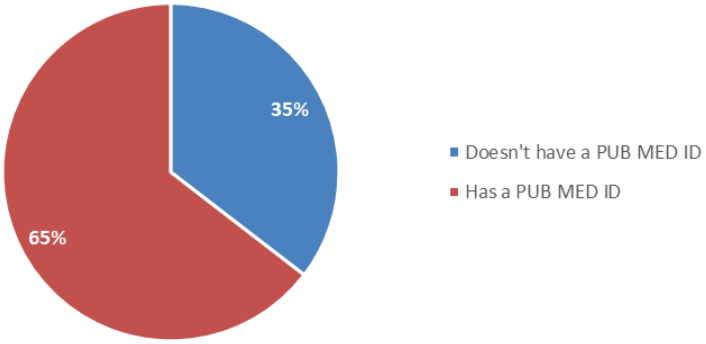


Table 2 shows the number of papers by the number of citations in Web of Science Core Collection as well as in all Web of Science databases. Most of the papers were not cited at all or received at most 9 citations. One paper was cited 935 times in the WOS Core Collection database and 956 times in all WOS databases.

Number of papers by number of citations; source: designed by the authors based on WOS data

Table 2

Times cite in WOS Core Collection	Number of papers	Times cite in all WOS databases	Number of papers
0	536	0	534
1 to 9	517	1 to 9	516
10 to 20	82	10 to 20	84
21 to 161	75	21 to 166	76
935	1	956	1

Table 3 presents the number of papers by university while Figure 5 displays the number of citations number of citations in WOS Core Collection per paper and the number of citations in all databases per paper by university. The highest number of papers is registered for the University of Medicine and Pharmacy "Carol Davila" Bucharest, yet the highest number of citations in WOS Core Collection as well as in all WOS Databases per paper is registered for the "APOLLONIA" University of Iași.

Number of papers by University; source: designed by the authors based on WOS data

Table 3

University	Number of papers
University of Medicine and Pharmacy "Carol Davila" Bucharest	298
"Grigore T. Popa" University of Medicine and Pharmacy	127
"Iuliu Haţeganu" University of Medicine and Pharmacy	126
Victor Babeş University of Medicine and Pharmacy Timişoara	117
Transilvania University of Braşov	98
"Lucian Blaga" University of Sibiu	67
Oradea University	65
"Dunărea de Jos" University of Galaţi	63
University of Medicine and Pharmacy Craiova	59
George Emil Palade University of Medicine, Pharmacy, Science, and Technology of Târgu Mureş	52
Ovidius University of Constanţa	45
Ştefan cel Mare University of Suceava	37
Titu Maiorescu University	24
University of Piteşti	11
Western University "Vasile Goldiş" from Arad	10
"APOLLONIA" University of Iaşi	7
"Constantin Brâncuşi" University of Târgu Jiu	5

Number of citations in WOS Core Collection per paper and Number of citations in all databases per paper by university; source: designed by the authors based on WOS data

Figure 5

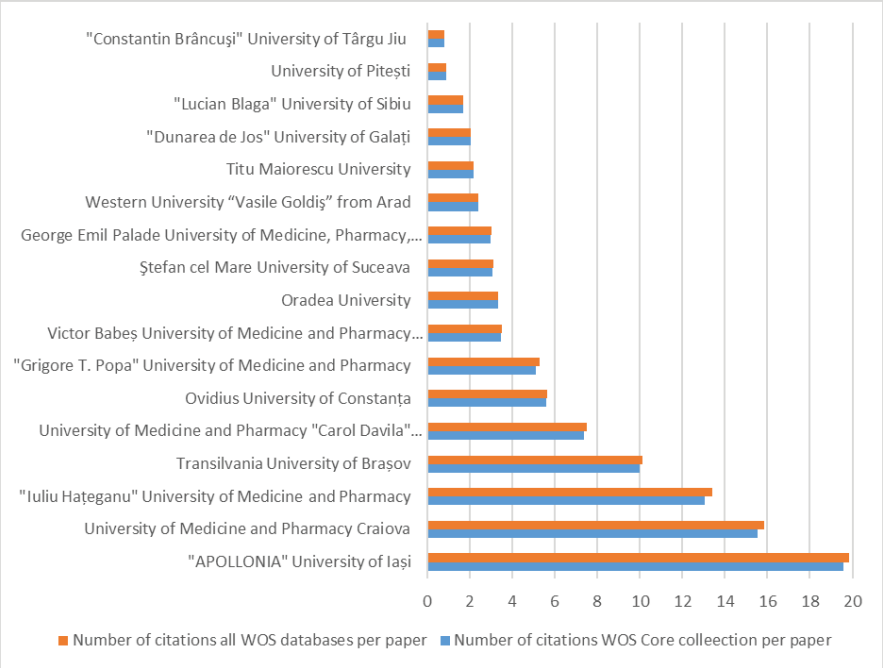


Table 4 shows the number of citations in WOS Core Collection and all WOS databases, the number of papers and number of citations per paper by university's property type. The same indicators by type of university (multidisciplinary or with medical study program only) are displayed in Table 5. Public universities have a higher number of citations per paper as well as a higher number of papers compared to private ones. Universities with only medical study program have approximately the same number of papers as multidisciplinary ones, yet a considerably higher number of citations per paper.

Number of citations in WOS Core Collection and all WOS databases, number of papers, number of citations per paper by university's property type; source: designed by the authors based on WOS data

Table 4

	Number of citations in WOS Core collection	Number of citations in all databases	Number of papers	Number of citations in WOS Core collection per paper	Number of citations in all WOS databases per paper
Private universities	213	215	41	5.20	5.24
Public universities	7782	7935	1170	6.65	6.78
Total	7995	8150	1211	6.60	6.73

Number of citations in WOS Core Collection and all databases, number of papers, number of citations per paper by type of university; source: designed by the authors based on WOS data

Table 5

	Number of citations in WOS Core collection	Number of citations in all databases	Number of papers	Number of citations in WOS Core collection per paper	Number of citations in all WOS databases per paper
Universities with only medical specialisations	5161	5265	600	8.60	8.78
Multidisciplinary university	2834	2885	611	4.64	4.72
Total	7995	8150	1211	6.60	6.73

Secondly, the results of the ANOVA procedure are presented (Table 6). There is a significant difference in the number of citations between papers with a PUB MED ID and papers without one. Similar results are obtained when the type of university is considered. However, there is no significant difference between papers published by affiliates in private universities or public ones. Yet, this might occur because there were only 41 of 1211 papers written by authors from private universities.

Results for the ANOVA procedure – Number of citations in WOS Core Collection and WOS all databases by several binary variables; source: designed by the authors based on WOS data

Table 6

	F-Statistic
Number of WOS Core Collection citations by whether or not the paper has a PUB MED ID	15.78*
Number of citations in all WOS databases by whether or not the paper has a PUB MED ID	15.78*
Number of WOS Core Collection citations by type of property of the university	0.06
Number of citations in all WOS databases by type of property of the university	0.09
Number of WOS Core Collection citations by type university (multidisciplinary or not)	4.94*
Number of citations in all WOS databases by type university (multidisciplinary or not)	4.93*

*Significant 5%

Tables 7 and 8 show the results of the logistic models 1 and 2. The Hosmer-Lemeshow test shows a probability above 5%, indicating a good fit of the model (Analyttica Datalab, 2019). Yet, the coefficients of X_4 and X_5 variables are significant. This shows that neither the type of university (multidisciplinary or exclusively medical) neither the property type of the university (public or private) are determinants of the likelihood that an article will be cited in WOS core collection or in all WOS databases. Tables 9 and 10 show the results of these models after eliminating non-significant variables. The results are very similar for models 1 and 2 because there are only 2 papers cited in other WOS databases but Core Collection. If the paper has a PUB MED ID and a higher number of parges, it has a higher chance to be cited. Also, recent papers have a lower probability to be cited.

Logistic regression results - model 1; source: designed by the authors based on WOS data

Table 7

Variable	Coefficient	z-statistic
X_1	1.87*	11.25
X_2	0.08*	7.41
X_3	-1.88*	-14.05
X_4	-0.20	-1.44
X_5	-0.32	-0.83
Prob (LR Statistic) = 0.00		
McFadden R-square = 0.24		
Hosmer Lemeshow Prob Chi-Square (8) = 0.47		

* Significant 5%

Logistic regression results - model 2; source: designed by the authors based on WOS data

Table 9

Variable	Coefficient	z-statistic
X_1	1.87*	11.25
X_2	0.08*	7.41
X_3	-1.88*	-14.05
X_4	-0.21	-1.44
X_5	-0.32	-0.83
Prob (LR Statistic) = 0.00		
McFadden R-square = 0.24		
Hosmer Lemeshow Prob Chi-Square (8) = 0.47		

* Significant 5%

Logistic regression results after removing non-significant variables – model 1; source: designed by the authors based on WOS data

Table 8

Variable	Coefficient	z-statistic
X_1	1.91*	11.25
X_2	0.08*	7.36
X_3	-1.89*	-14.08
Prob (LR Statistic) = 0.00		
McFadden R-square = 0.24		
Hosmer Lemeshow Prob Chi-Square (8) = 0.17		

* Significant 5%

Logistic regression results after removing non-significant variables – model 2; source: designed by the authors based on WOS data

Table 10

Variable	Coefficient	z-statistic
X_1	1.92	
X_2	0.08	
X_3	-1.88	
Prob (LR Statistic) = 0.00		
McFadden R-square = 0.24		
Hosmer Lemeshow Prob Chi-Square (8) = 0.21		

* Significant 5%

4. CONCLUSIONS

The rapid publication of related literature for the new disease is an essential and valuable resource for the support of the medical and scientific research community. Although, accelerating the speed of dissemination in response to the worldwide phenomenon identified as COVID-19 pandemic highlight significant changes in characteristics of scientific research published in medical journals, it must not compromise quality and ethical standards.

Our research aimed to assess what is the impact of COVID-19 papers written by authors affiliated to Romanian universities with a medicine specialisation, which are the factors that influence the probability of a paper being cited and how does the number of citations vary by university type.

The study analysed data from a total of 18 Universities from Romania, both public and private. We identified that half of the papers have between 1 and 10 pages long while 36% of them are 11 to 20 pages long, with approximately half of all being published in journals indexed Science Citation Index Expanded. The results also reveal that the highest number of published papers is registered for the University of Medicine and Pharmacy "Carol Davila" Bucharest, although "APOLLONIA" University of Iași has the highest number of citations in WOS Core Collection as well as in all WOS Databases per research.

The ANOVA procedure outputs reveal significant difference in the number of citations between papers with a PUB MED ID and papers without one. Moreover, the results of our logistic models shows that neither the type of university (multidisciplinary or exclusively medical), neither being of public or private type are determinants of the probability for an article to be cited in WOS core collection or in all WOS databases. Consequently, a paper with PUB MED ID and a higher number of pages, has a higher chance to be cited.

In future work, we plan to extend our research to track the existence of gaps within and between groups in terms of article publishing in Romania.

5. References

1. **Garrett, L.** (2020). *COVID-19: The medium is the message*. Lancet, 395, 942-943.
2. **Limaye RJ, Sauer M, Ali J, Bernstein J, Wahl B, Barnhill A**, Labrique A. Building trust while influencing online COVID-19 content in the social media world. Lancet Digital Health. 2020;2(6):e277–e278.
3. **Horbach SPJM**. (2020) *Pandemic publishing: Medical journals strongly speed up their publication process for COVID-19*. Quant Sci Stud. 2020;1: 1056–1067.
4. **Bagdasarian, N., Cross, G.B. & Fisher, D**. *Rapid publications risk the integrity of science in the era of COVID-19*. BMC Med 18, 192 (2020).
5. data.gov.ro, Open datasets portal, https://data.gov.ro/dataset/studenti-inmatriculati-la-forma-de-studiu-licenta-la-1-ianuarie-2020, accessed April 30th 2022.

6. Analyttica Datalab (2019) https://medium.com/@analyttica/hosmer-lemeshow-goodness-of-fit-test-65b339477210, accessed April 30th 2022.
7. **Andrei, CL., Catană, A. et al**. (2021) *The impact of Covid-19 lockdown on the level and intensity of physical activity: a statistical analysis on Romanian young adults*, The Medical-Surgical Journal (Rev Med Chir Soc Med Nat, Iaşi), Vol. 125, Nr. 4, pg. 550-562, 2021.
8. **Paun, M. et al**. (2020) *Predicting long-term citation counts in Web of Science: COVID-19 early publications case study*, Romanian Statistical Review nr. 4, 65-80, 2020.
9. **Albu, L.L. et al**. (2020) *Estimates of Dynamics of the Covid-19 Pandemic and of its Impact on the Economy*. Rom. J. Econ. Forecast. 2020, 13, 5–17.
10. **Păunescu, C., & Matyus, E**. (2020) *Resilience measures to dealing with the COVID-19 pandemic. Evidence from Romanian micro and small enterprises.* Management & Marketing-Challenges for the Knowledge Society, 15, 439–457, 2020.
11. **Onofrei M. et al.** (2021) *Effects of the COVID-19 pandemic on the budgetary mechanism established to cover public health expenditure. a case study of Romania*. Int J Environ Res Public Health 18(3):1134, 2021.
12. **Raiu, C., & Mina-Raiu**, **L.** (2022) *How to Cope with Counter-Performance in Public Administration. The Case of Freedom of Religion or Belief During the Pandemic.* Transylvanian Review Of Administrative Sciences, 18(66), 81-98. 2022.
13. **Raiu, C. V., & Juknevičienė, V**. (2021) *The Science of the State. A comparative analysis of Statistical National Offi ces in the European Union.* Romanian Statistical Review, (2) 14-28. 2021.
14. **Ionescu, C.A.** et al. *Sustainability Analysis of the E-Learning Education System during Pandemic Period—COVID-19 in Romania*. Sustainability 2020, 12, 9030.
15. **United Nations**. *Transforming our world: the 2030 Agenda for Sustainable Development* A/RES/70/1, 2015.

Job offer on the Romanian labor market in digital economy

Carmen-Elena BĂNESCU (banescu_carmen96@yahoo.com)
Bucharest University of Economic Studies, Bucharest, Romania

Emilia ȚIȚAN (emilia.titan@csie.ase.ro)
Bucharest University of Economic Studies, Bucharest, Romania, Institute on National Economy

Daniela-Ioana MANEA (daniela.manea@csie.ase.ro)
Bucharest University of Economic Studies, Bucharest, Romaniay, Institute on National Economy

ABSTRACT

The Fourth Industrial Revolution had a major impact on the labor market. There are many consequences of the digitalization of the economy, but this research shows the changes in supply structure on the Romanian labor market. The companies that are operating in Romania usually publish job opportunities on recruitment platforms. Using the web scraping techniques we will identify which jobs are currently available on the Romanian labor market, in which cities, but also the field of activity. First of all, we will analyze the concentration of jobs in cities. We will classify these jobs based on an official classification code used by the Romanian state, because the job offer is very wide. Moreover, we will analyze the level of experience and the level of training that the candidate must have.

In Romania, the main economic centers are Bucharest, Timisoara and Cluj-Napoca. In these three cities we also identify the highest concentration of jobs. There are many jobs in the IT field and those jobs that involve communication. Candidates with at least one year of experience in the field are preferred, as well as candidates who are willing to work full time. It can be noted that jobs with repetitive activity have almost disappeared, because they have been replaced by industrial robots. Also, it can be seen that employers are looking for candidates with higher education. In this step we will talk about job polarization phenomenon.

Finally, the paper aims to indicate the stage of development of the Romanian labor market and what young people need to know before choosing their future career. Also, this paper is the starting point of a wide research about digitalization on labor market.

Keywords: *job offer, online platforms, digital economy, text mining, text processing, web scrapping, job polarization*

1. INTRODUCTION

Today's society is facing the Fourth Industrial Revolution, which is based on the automation of industrial processes and on the use of hardware equipment to control them and an unprecedented development of artificial intelligence. As a result, all industries need to adapt in order to survive. The digitalisation of the economy has a major impact on the labor market, especially in less developed countries. Moreover, one of the effects of the digitalization of the economy on the supply of jobs is their polarization.

The topic addressed in the research will focus on the supply of jobs on the Romanian labor market and whether the polarization effect determined by the digital economy is observed. First of all, the specialized literature will be presented, in order to describe the implications that the digitalization of the economy has on the supply on the labor market. This section is followed by a presentation of the changes that have occurred in the job offer over the years in Romania. More specifically, the evolution of the job offer for the main groups of occupations (according to the Code of Classifications of Occupations in Romania) in recent years will be highlighted, based on the official reports of the National Institute of Statistics. In addition to the above, the jobs currently available on online recruitment platforms will be analyzed. The main groups of occupations of interest for companies, the level of training required of the candidates, but also the main skills and knowledge that the candidate must have will be presented. Finally, the development stage of the Romanian labor market and what young people need to know before choosing their future career will be presented.

2. LITERATURE REVIEW

Even though the phenomenon of digitization has been going on for decades, the development of Big Data and robotics heralds a new economy (Degryse, 2016). The impact of digitalisation on the labor market has become a rather controversial topic and has been discussed for the last 30 years. This is largely due to the complexity of the relationship between digitization and jobs (Valenduc & Vendramin, 2016). The development of technology is seen as an important factor in increasing the number of unemployed, both in Romania and in Western countries. The changes in the labor market caused by digitalisation are not necessarily negative. Like any invention, the discoveries in technology have as their main purpose the evolution of society. The automation of some processes will determine the disappearance of some jobs, but also the appearance of others. This aspect can rather be considered a structural change

in employment according to industry, occupation, qualification and tasks (Ulrich, 2016). However, the risk of automating certain jobs depends very much on the cost of developing the solution, the dynamics of the economy and the economic benefit of this solution (Heald, Smith, & Fouarge, 2019).

Companies active in the market that have had investments in the direction of digitalization have started to reconfigure their business model trying to stay competitive. With this reconfiguration comes the need for specialized employees. Thus, the labor market supply focuses rather on jobs that require a high and low qualification, which do not involve repetitive activities. This trend has been called "job polarization". This phenomenon means an increase in the employment rate of well-paid specialists and managers, but also of low-paid workers in services and a reduction in the demand for workers in the manufacturing industry and employees performing repetitive activities (Goos, Manning, & Salomons, 2014).

In April 2020, a study was published analyzing the polarization phenomenon in Central and Eastern Europe (CEE). The conclusions of the study confirm the hypothesis that the demand for jobs that require an average level of training decreased by about 9 percentage points in 2016, compared to 2000. Moreover, research shows that specialists are the most sought after in the labor market work in the CEE, identifying an increase of about 69% in 2016, compared to the year 2000 (Nchor & Rozmahel, 2020).

Regarding the skills of Romanians, the European Center for the Development of Vocational Training has built an index that measures the performance of skills systems at European level. In this research, Romania was ranked in the group of countries with poor results. Romania has the weakest performance in skills development due to lack of access to technology and training. The most undesirable aspect is that 15-year-old students fail to get good scores in terms of reading, applying mathematics or science knowledge. Romania also occupies a low position in the number of graduates who managed to get hired. There is also a relatively positive aspect, namely the matching of skills with the requirements of the job (European Centre for the Develeopment of Vocational Training, 2020).

Moreover, the European Center for the Development of Vocational Training also analyzed the job offer published online in Romania during 2018-2019. The study shows that 47.4% of jobs require employees the ability to adapt to change, while the use of communication techniques is the least common ability (European Centre for the Develeopment of Vocational Training, 2020).

3. METHODOLOGY

The proposed case study refers to the analysis of jobs published on recruitment platforms in Romania. In this approach, the following steps will be completed: data description, data processing and subsequently data analysis.

The data used in the case study were obtained through the web scraping technique from an online recruitment platform widely used in Romania. Basically I built a recursive script in Rstudio that extracts certain objects from the HTML page. This automatic data extraction technique is possible due to the HTML language used to create web pages and which is a structured language (Landers, Cavanaugh, Brusso, & Collmus, 2016). Thus, in October of this year on this online platform were active about 2400 jobs and trainings published by companies operating in our country. From the web pages of the platform we selected the name of the offered position, the field of activity, the location, the required level of experience, the availability necessary for this job, but also the description of the offered position. The jobs published on these recruitment platforms do not detail all the requirements that the future candidate must meet. Moreover, many of the jobs do not actually present the activity or requirements, but the company that offers the job. Because of this, it was necessary to map jobs based on the Romanian Occupational Classifications Code.

In the data processing stage we removed punctuation marks, special characters, numbers and connecting English words from job descriptions. Then we applied tokenization method, which means dividing sentences into words. After that, the words are lemmatized. Lemmatization method means transforming words into basic form. The processed corpus will be transformed into the Term Document Matrix. This matrix will be used to apply text mining methods.

The analysis of texts written in natural language can be considered a machine learning technique. It is not a fixed method, it can be customized depending on the topic approached and the answers you want to get. Text analysis starts from calculating frequencies, classification, recognition of entities, identification of topics, sentiment analysis and many other derivatives.

In the case study I will analyze an unstructured data collection using Topic Modeling, a text mining technique that establishes semantic patterns. This statistical method is a form of unsupervised classification. The natural language processing method is known as Latent Dirichlet Allocation or LDA. The mathematical model is based on two major principles. Each data collection contains a variety of topics covered and in turn these topics are composed of several words of specific importance. Basically, the Latent Dirichlet Allocation

method in this case will identify the collection of words that define a group of characteristics, while determining the characteristic groups that define the entire job base.

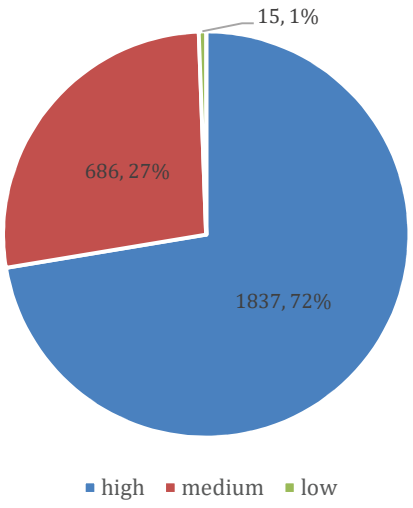
The whole process of data collection, data processing and data analysis was done with the help of RStudio software. For data collection we used functions of rvest, plyr, xml2 packages. In the data processing stage we used functions from the tidyverse and tm packages offered by Rstudio. In the data analysis stage we used both the topicmodels package and other packages designed to obtain graphics, such as wordcloud and shiny. Additionally, we used simple graphs built in Microsoft Office Excel.

4. RESULTS AND DISCUSSION

According to the National Institute of Statistics, the number of vacancies increased considerably between 2012-2019. The year 2020 is more affected in this respect, due to the global pandemic and the crisis it has caused in the economy. The results show that the supply of active jobs available on the market in October indicates a higher demand for jobs that involve higher training.

Job distribution by level of studies

Figure no. 1

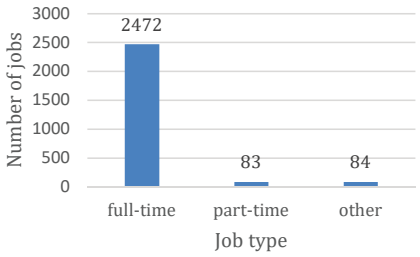


Source: Generated by author in Microsoft Excel and using data from online platform

From the 2,538 positions, approximately 72% correspond to a candidate profile with higher education, 27% of the jobs are for people with secondary education, and the remaining 1% are jobs that require low-level education. Employers want the most specialists in various fields of activity in proportion of approximately 66%, but also technicians and other specialists in the technical field in proportion of 16%. Administrative officials, service workers, skilled and assimilated workers, operators of installations and machines, assemblers of machines and equipment, but also elementary occupations reach a percentage of 12% of the total jobs posted.

Job distribution by job type

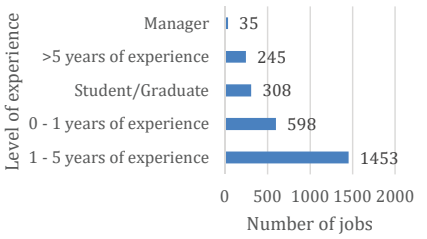
Figure no. 2



Source: Generated by author in Microsoft Excel and using data from online platform

Job distribution by level of experience

Figure no. 3

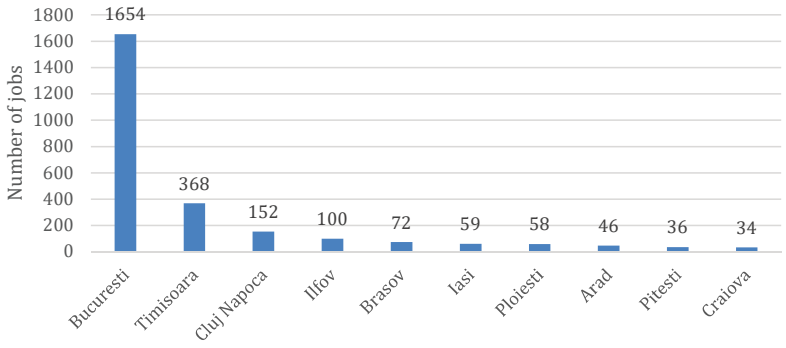


Source: Generated by author in Microsoft Excel and using data from online platform

In terms of the type of job, the largest offer is for people who are willing to work full time and for those who have a level of experience at least 1 year in the field.

Job distribution in top 10 Romanian locations

Figure no. 4



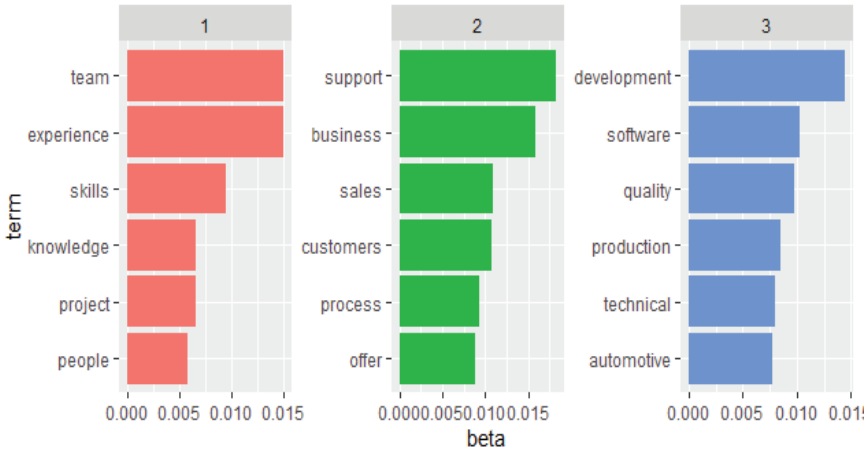
Source: Generated by author in Microsoft Excel and using data from online platform

A very important aspect to mention is the country-wide distribution of these jobs. It seems that the jobs offered by employers on online recruitment platforms are mainly from the big economic centers in Romania. This is normal, as digitization is not widespread in all regions of the country.

In addition to the quantitative aspects extracted immediately from the database, from the processing of the job descriptions, we extracted the main topics addressed in them. There are jobs for which there is no description, which is why they have been excluded in order to obtain reasonable and relevant results. A fairly important parameter in the Latent Dirichlet Allocation model is the beta parameter, which helps to rank words within topics. Often, due to the texts used, the results are more or less strong.

Three topics discovered and its keywords

Figure no. 5



Source: Generated by author in RStudio and using data from online platform

We considered of interest three topics that we renamed according to their scope: experience, soft skills and technical skills. The model assigns to each record the list of the three topics, but also a gamma parameter, which helps us decide which topic is right for the description of that job. Thus, 49% of the jobs require some work experience in the field, for 28% of the jobs soft skills are important, and for 24% of the jobs analyzed the technical skills are a priority.

Latent Dirichlet Allocation, however, did not highlight the topics as desired, which is why we completed the analysis with a simple analysis to extract some patterns from the data indicating exactly what was observed in the previous method, but not as clearly.

Soft and technical skills

Table no. 1

Soft skills	%Rate
Relevant experience	55,92
Foreign languages	41,27
Communication skills	35,16
Team work	14,52
Analytical	13,50
Problem solving	11,08
Motivation	8,66
Creativity	7,64
Negotiation skills	3,18
Time management	1,91
Adaptability	1,53
Interpersonal skills	0,76
Open mind	0,76

Technical skills	%Rate
Programing software	23,82
Microsoft office	18,09
Accounting	21,91
Data science	12,23

Source: Generated by author in Microsoft Excel and using data from online platform

The results obtained from the analysis of the data obtained for a short time window are somewhat similar to the analysis developed by the European Center for the Development of Vocational Training, which identifies as essential skills those specified above.

The most common skills required

Figure no. 6

Source: Generated by author in RStudio and using data from online platform

It can be seen that in general, the jobs on the recruitment platforms want the future candidates to have very well defined so-called soft skills. Obviously, the technical knowledge is not left to chance, but they are better

focused in terms of jobs dedicated to specialists. As a result of the digitization and automation of repetitive processes, the skills that these routine jobs required are no longer needed. It is very important for employers to ask candidates for skills that will train their thinking and communication skills. Candidates must be able to make decisions in certain new situations. The digital age is actually about the ability of people to create, to imagine solutions or how technology could help them in their workplace. Obviously, it is important that employees also have technical skills related to technology, because many activities become easier.

It is known that large corporations want to produce more and become more profitable, so they invest in technology and people who know how to cope in such an environment. In fact, jobs are not disappearing, but structurally changing.

Some employees do not adapt, which is why unemployment occurs. There are also employees who have speculated on the benefits of the changes and that their work has become easier.

It is not enough for only companies to want this change, but also for the state. Usually, if a company does not have enough human capital adapted to the new needs of the company, it decides to leave that country. The human resource adapted to the new economic era must be prepared through an updated educational system. This approach is quite complicated in Romania, because in some regions of the country students do not have access to the minimum education. It is true that these shortcomings cannot be radically changed, because they depend on many economic and social factors about which I do not have much information. It can also start with a closer collaboration between state bodies and companies operating on the Romanian market. There are companies that see a good potential in the Romanian human resource and are unconditionally involved, but there is also the option for the state to offer facilities to companies to intervene in the university training process and not only.

As we have noticed, specialists are the main attraction on the Romanian labor market. Most of the jobs extracted from the database are in the field of IT Hardware and IT Software. As a result, it is interesting to analyze which are the most wanted features, as they would have more homogeneity in terminology and would lead to a more consistent analysis

Romanian 3-digit Code of Occupations Classifications

Table no. 2

Romanian 3-digit Code of Occupations Classifications	Number of jobs
Software analysts	561
Finance specialists	290
Specialists in administrative field	239
Specialists in sales, marketing and public relations	230
Agents and brokers in sales and supply	138
Leaders in sales, marketing and development	118
Engineers	117
Sellers in stores	109
Technicians for operations in the field of information and communication technology and for providing technical support for users	82
Specialists in databases and networks	77

Source: Generated by author in Microsoft Excel and using data from online platform

It is observed that in the top 10 jobs according to the Romanian 3-digit Code of Occupations Classifications, software analysts in the software field are the most sought after. As previously mentioned, we have also accumulated specialists in databases and networks, because they are part of the IT field.

The most common skills required in IT field

Figure no. 7

Source: Generated by author in Rstudio and using data from online platform

It can be seen that the most wanted skill is problem solving, even if the jobs are technical. This is not an unexpected aspect, as the creation of algorithms relies heavily on the developer's analytical ability. I say this

because artificial intelligence has become so advanced that it will write its own code, but it is necessary for a specialist to take care of maintenance or change direction when necessary.

The main skills identified in IT jobs are both soft skills (relevant experience-25.44%, communication skills -12.63%, foreign languages -10.53%, problem solving -8.25%, creativity-7.54%, motivation-6.49%, team work-4.04%, analytical thinking-2.63%, adaptability-0.35%) and technical skills. Technical skills mean both programming languages (C++ -5.79%, cloud-5.09%, java-3.86%, R -3.16%, .NET-1.58%, Angular-1.58%, Python-1.58%, Amazon web services-0.88%, CSS-0.53%, ETL-0.35%, MatLab-0.35%, Network protocols-0.35%, Perl-0.35%, Ruby-0.35%, Cobol-0.18%, Firewall-0.18%, Kafka-0.18%, Linux-0.18%, Scala-0.18%, Wide Area Network-0.18%, Hadoop-0.18%), programming concepts (Scalable programming-1.05%, Object Oriented Programming -0.7%, Optimization-0.7%,) and databases (SQL/NoSQL -3.8%), as well as theoretical and practical knowledge in the area of data science (Computer science -4.21%, Big Data-0.7%, Forecasting-0.35%, Clasification-0.18%, Machine Learning-0.18%, Regression-0.18%, Data visualization-0.18%).

Previously, we presented the most known and most wanted soft skills, technologies, programming languages and technical knowledge that IT specialists need to understand and know. Obviously, there are many technologies to which reference is made, but these are the most appreciated and with a fairly wide spectrum of use in any field of activity.

5. CONCLUSION

The labor market in Romania, referring to the supply of jobs on online platforms indicates a strong polarization. This aspect confirms that the European trend demonstrated by other researchers is also maintained in Romania. There is a predominant requirement for specialists and less for mid-level employees. The shortcoming of online platforms in Romania is that there are no restrictions when the employer submits the job. Most of the time, instead of the description, nothing is found or a general description of the company is found, which is not exactly right. Because of this, although the number of jobs is reasonable, textual modeling is not very effective, as texts are often lacking in information. In the current epidermiological context, these platforms are very useful, but the candidate loses a lot of time only to identify if that job is suitable for his training. As time is one of the most important resources, a more detailed specification of the main requirements and benefits that the company has would exponentially reduce the recruitment process.

However, the information provided was sufficient to outline the job offer on the Romanian market. One of the most important strengths required by employers is experience, but often young people do not have this experience. Therefore, recruitment companies periodically organize various events, trainings, workshops, to which top companies operating on the market are invited. Thus, young people have the opportunity to try new things and much closer to what employers want. Moreover, it is important that when young people choose a field of study to focus on the three aspects described above: experiences, soft skills and technical skills. They certainly cannot decide for themselves which direction they should go, but here a career counseling organized by the educational institutions themselves should intervene.

References

1. **Degryse, C**. (2016). *Digitalisation of the economy and its impact on labour markets.* European Trade Union Institute.
2. **European Centre for the Develeopment of Vocational Training**. (2020, Octomber 1). *European Skills Index*. Retrieved from Skills Panorama: https://skillspanorama.cedefop.europa.eu/en/dashboard/european-skills-index?year=2020&pillar=&country=RO#3
3. **European Centre for the Develeopment of Vocational Training**. (2020, Octomber 1). *Skills in online job advertisements*. Retrieved from Skills Panorama: https://skillspanorama.cedefop.europa.eu/en/dashboard/skills-online-job-advertisements?country=RO&occupation=#29
4. **Goos, M., Manning, A., & Salomons, A**. (2014). *Explaining job polarization: routine-biased technological change and offshoring*. American Economic Review, 2509-2526.
5. **Heald, S., Smith, A., & Fouarge, D**. (2019). *Labour market forecasting scenario's for automation risks: Approach and outcomes.* Cedefop Skills Forecast 2018.
6. **Landers, R. N., Cavanaugh, K. J., Brusso, R. C., & Collmus, A. B**. (2016). *A Primer on Theory-Driven Web Scraping: Automatic Extraction of Big Data from the Internet for Use in Psychological Research*. Psychological Methods, 475-492.
7. **Nchor, D., & Rozmahel, P.** (2020). *Job Polarization in Europe: Evidence from Central and Eastern European Countries*. Law, Economics and Social Issues Review, 52-74.
8. **Ulrich, W**. (2016). *Digitalization and structural labour market problems: Tha case of Germany*. International Labour Office, 17.
9. **Valenduc, G., & Vendramin, P.** (2016). *Work in the digital economy:sorting the old from the new*. European Trade Union Institute.

Increasing RDI Outputs through the Competitive Research Funding Operational Programme with Impact on the Emerging Market

Carmen DOBROTA
Faculty of Administration and Business, University of Bucharest

Maria-Magdalena ROSU
Faculty of Cybernetics, Statistics and Informatics, The Bucharest Academy of Economic Studies

Andreea-Ionela PUIU
Faculty of Cybernetics, Statistics and Informatics, The Bucharest Academy of Economic Studies

Eduard C. MILEA
Faculty of Cybernetics, Statistics and Informatics, The Bucharest Academy of Economic Studies

Mihaela PAUN[1] (mihaela.paun@faa.unibuc.ro)
Faculty of Administration and Business, University of Bucharest,
Research Institute of the University of Bucharest–ICUB, University of Bucharest
National Institute for Research and Development for Biological Sciences

ABSTRACT

Considering the broad impact of applied research on the economy, RDI funding evaluations are required both in terms of the number of allocated resources and the management of these resources. RDI efficacy depends on human resources productivity and the financing instruments established through national priorities. This paper offers an analysis of the RDI investments in Romania allocated through the Competitiveness Operational Programme 2014-2020, covering the European Structural and Investment Funds, namely the European regional development fund. The analysis of the funded projects highlighted the common trends among the beneficiaries of RDI projects, their options in managing resources in relation to the eligibility of costs, and their national distribution between the seven development regions of Romania. The amount of funding was discussed in terms of the smart specialization domains es-

1. Corresponding author

tablished by the national strategy. The conclusions of the study, correlated with the objectives set by the SNCDI 2014-2020, are relevant for the management of the future funding instruments allocated to RDI by the ERDF in the period 2021-2027.

Keywords*: competitive research funding, RDI funding evaluations, financing instruments, national strategy, funding eligibility*

JEL Classification: Z18

1. INTRODUCTION

The European Regional Development Fund (ERDF) aspires to strengthen the social and economic cohesion in the European Union. The ERDF funding programs are governed by Regulation (EU) No. 1303/2013 and Regulation (EU) No. 1301/2013. While the former regulation lays down the common rules applicable to the ERDF, the latter encompasses specific provisions concerning the Investment for growth and jobs goal.

Each Member State elaborates a Partnership Agreement (PA) covering all Programmes of the European Structural and Investment Funds (ESIF), including the European regional development fund (ERDF). Considering the funding priorities set out in the PA, Romania has implemented the Competitiveness Operational Programme 2014-2020 (Autoritatea de Management pentru Programul Operațional Competitivitate, n.d.).

The PA identifies five main challenges for the development of Romania, with competitiveness being specifically highlighted. The agreement underlines the need to improve the innovation capacity for developing new products and services, to improve the business environment through the implementation of scalable value chains. Additionally, COP contributes to fulfilling the objectives of three other growth challenges: infrastructure, people and society, administration and governance. This enables horizontal interventions in the economy and society.

Increasing investments and stimulating research, development and innovation (RDI) activities can increase competitiveness through smart specialization, added value, and internationalization. One of the objectives of the Europe 2020 Strategy is to reach a level of 3% of EU GDP for research and development expenditure. Romania has assumed a target of 2% of GDP for RDI financing (1% of GDP public expenditure and 1% of GDP expenditure from private sources). This is a very ambitious target given that in 2011 Romania invested only 0.48% of GDP in RDI, 80% of investments being made by the public sector (National Strategy for Research, Development and Innovation 2014-2020, 2018) and it has decreased considerably during the 2014-2020 implementation period. Towards the goal of achieving research investment targets, prerequisites are put in place through a mix of measures and policies

designed to prioritize and focus public resources, on one hand, and encourage private investment and technological development, on the other hand.

1.1. The Competitiveness Operational Program in Romania

The Competitiveness Operational Programme 2014-2020 divides its mission to increase competitiveness in two specific directions:

1. Priority Axis 1 (PA1) - Research, development, and innovation in support of economic competitiveness and business development.

2. Priority Axis 2 (PA2) - Information and Communication Technology.

For the purpose of this paper, we will focus solely on PA1, which addresses market needs, such as creating a more compact and modern RDI environment, or creating an entrepreneurial and innovative culture in RDI, both in the public and the private sector, that contributes to training and development of latent potential in the field. The three types of thematic priorities identified by NSRDI on which PA1 focuses are the following:

- Smart specialization priorities - defining and consolidating areas of high competence where there are real or potential competitive advantages and which by directing resources and organizing a critical mass of researchers, can ensure competitiveness in regional and/or global value chains.
- Priorities with public relevance - developing the capacity of the public sector to scan the space of new and promising technologies and to request novel solutions from public and private RDI operators.
- Fundamental research (funded mainly from NPRDI III and the programs of the Romanian Academy).

Romania invests only 0.3% of GDP in research and development, compared to the 1% assumed target, placing it in the least favourable place in the European Union and making it difficult to effectively stimulate competitiveness.

Nationally, the COP interim evaluation underlines an unbalanced effective execution of funds among regions, with the Bucharest-Ilfov region having contributed almost two-thirds to the achievement of the total expenses from the activity of R&D. On the other hand, the most significant increase compared to 2014 was registered in the West region and the smallest in the North-East region. In addition, although the number of units with research and development activity decreased during this period, there is a slight increase in the number of employees in research and development, mainly due to the positive evolution of the business environment (Autoritatea de Management pentru Programul Operațional Competitivitate, n.d.).

After the 1988 reform of the Structural Funds (Bachtler and Michie, 1995) the systematic evaluation of EU regional policy became proprietary. An ex-post evaluation of a program at the end of its implementation period will provide a clear picture of its effectiveness and insights for the next funding period.

The current study proposes a descriptive analysis of the "Competitiveness Operational Programme 2014-2020" implementation. Based on the aforementioned strategies to increase competitiveness through the consolidation of areas of high competence, we considered the following evaluation objectives:

Objective 1: To identify the common tendencies among fund recipients.

Objective 2: To identify the regional competencies.

Two more evaluation objectives tackle the need for financial incentives to increase competitiveness:

Objective 3: A description of the non-eligible value of the projects.

Objective 4: A description of eligible salaries and non-reimbursement salaries tendencies.

2. METHODS AND MATERIALS

The dataset was obtained from the Ministry of Research, Innovation and Digitalization, under the provisions of the Law no. 544/2001 regarding free access to public information. Statistical analysis was done through R statistical software, version 4.1.1, while the graphics were done in Exploratory version 6.6.3

Data on funded projects under the Priority Axis 1 of the Competitiveness Operational Programme 2014-2020 consists of 331 projects classified in five different types and seven actions namely, Action 1.1.1 Large R&D infrastructures (14.5%), Action 1.1.2 Development of networks for R&D centres (3.6%), Action 1.1.3 Synergies with Horizon 2020 (8.5%), Action 1.1.4 Attract researchers from abroad (15.4%), Action 1.2.1 Stimulate the demand of enterprises for innovation (41.7%), and Action 1.2.3 Knowledge transfer partnerships (16.3%).

The National Strategy for Research, Development and Innovation (NSRDI) defined the five areas of smart specialization for the strategic cycle 2014-2020, based on their scientific and commercial potential. The Competitiveness Operational Programme PA1 had the following smart specialization project distribution: Bioeconomy (12.8%), IT&C, space and security (25.8%), Energy, environment and climate change (16.9%), Eco-Nano-Technologies and advanced materials (27.9%) and Health (16.6%).

Eligible for funding and implementing projects under the Competitiveness Operational Programme are various organizations, classified by their legal status in ten categories. Based on the evaluation process, the distribution of funded projects implemented by each type of institution is as follows: National Institutes for R&D (NIRD) (17.5%), Universities (25.1%), Research Institutes of the Romanian Academy (RIRA) (3.9%), Microenterprises (41.4%), Medium enterprises (4.8%), Large enterprises (3.0%), Associations (0.6%), Foundations (0.9%), Hospitals (1.2%) and another type of beneficiaries (public institutions with research activities; 1.6%).

According to the provisions of EU regulations for the period 2014 - 2020, Romania's development regions fall into two categories: more developed regions (Bucharest - Ilfov Region, which includes the capital Bucharest) and less developed regions (the other 7 development regions of Romania, respectively North-East, South-East, South Muntenia, South West Oltenia, West, North-West and Center).

The allocation of projects for each region is as follows: Bucharest - Ilfov (48%), Center (6.3%), North East (8.8%), North West (16.6%), South East (2.4%), South Muntenia (4.8%), South West Oltenia (7.6%) and West (5.5%).

In the Competitiveness Operational Programme PA1, the Bucharest - Ilfov Region will benefit from a total EU contribution of 249 million euros, while the other seven regions combined will benefit from a total EU contribution of 1.08 billion euros (the distribution of the EU contribution takes into account the project's location of implementation).

The amount (in thousands of euros) of funded projects and eligible and non-eligible contributions is represented in Table 1.

Project funding metrics

Table 1

	Min	Max	Mean	Median	SD
Total value of the project (in thousands of euros)	54.19	157574.81	2907.50	1405.93	9570.13
Eligible value of the project (in thousands of euros)	51.80	157574.81	2670.19	1285.96	9417.79
Non eligible value of the project (in thousands of euros)	0	8613.84	238.03	17.63	798.84
Eligible salaries (in thousands of euros)	13.39	3327.95	624.52	412.07	623.05
Non-reimbursement salaries (in thousands of euros)	3.68	3327.95	558.57	376.29	541.44

Objective 1: Common tendencies among funds recipients

Investigating the common trends among beneficiaries, the following has been identified.

Action 1.1.1 - Large R&D infrastructures.

- Section A (Investments for enterprises R&D departments) - 22 projects were funded, with a total value of 124.96 million euro (out of which 44.20 million euro in the European Regional Development Fund (ERDF)).
- Section B (2015 – Innovation clusters) - 4 projects were funded, with a total value of 10.21 million euro (out of which 4.92 million euro in ERDF).
- Section F (Investment projects for public R&D institutions/ universities) - 18 projects were funded, with a total value of 2.04 billion euro (out of which 4.92 million euro in ERDF).
- "Extreme Light Infrastructure - Nuclear Physics (ELI-NP)" - one project was funded with a total value of 157.87 million euro (out of which 126.98 million euro in ERDF). The fund recipient for this project was the national institute for research and development implementing this project.
- "DANUBIUS-RI" - one project was funded with a total value of 4.52 million euro (out of which 3.84 million euro in ERDF). The fund recipient for this project was the national institute for research and development implementing this project.

Overall, for this action, the national institutes of research and development were the most active in securing funds for Section F projects. Among the five domains from the NSRDI, Eco-Nano-Technologies and advanced materials and Health were the ones with the largest share in the COP funding. The recipients with the largest percentage of funds were the national institutes of research and development beneficiaries.

Action 1.1.2 - Development of networks for R&D centres

- Section "Cloud type projects and massive data infrastructures" - 11 projects were funded, with a total value of 10.87 million euro (out of which 8.87 million euro in ERDF).

For this action the largest number of beneficiaries were the universities, implementing projects under the IT&C, space and security - smart specialization domains and Health as priority domain of national interest, however the highest percentage of contracted values was secured by ONG's beneficiaries.

Action 1.1.3 - Synergies with Horizon 2020

- Section "RO ECSEL" - 11 projects were funded, with a total value of 11.32 million euro (out of which RON 8.33 million euro in ERDF).
- Section "Complement" - 3 projects were funded, with a total value of 5.02 million euro (out of which 4 million euro in ERDF).
- Section "Support Centre" - 13 projects were funded, with a total value of 7.32 million euro (out of which 1.08 million euro in ERDF).

The largest numbers of beneficiaries in this action were the universities, implementing projects under the IT&C, space and security and the Eco-Nano-Technologies and advanced materials smart specialization domains. The highest percentage of contracted values was secured also by university type beneficiaries.

Action 1.1.4 - Attract researchers from abroad

- Section E (Attracting staff with advanced skills from abroad to strengthen R&D capacity) 51 projects were funded, with a total value of 86.39 million euro (out of which 69.13 million euro in ERDF).

The most salient beneficiaries in this action were the universities, implementing projects under the Eco-Nano-Technologies and advanced materials as smart specialization domains and Health as priority domain of national interest. The highest percentage of contracted values was secured also by university type beneficiaries

Action 1.2.1 - Stimulate the demand of enterprises for innovation

- Section C (Innovative start-up and spin-off enterprises) - 91 projects were funded, with a total value of 17.82 million euro (out of which 12.21 million euro in ERDF).
- Section D (New innovative start-ups) - 22 projects were funded, with a total value of 24.27 million euro (out of which 17.82 million euro in ERDF). All funds were assigned to small and microenterprises.
- Section "Innovative technological project" - 25 projects were funded, with a total value of 24.27 million euro (out of which 17.81 million euro in ERDF).

The most salient beneficiaries in this action were microenterprises implementing projects under the Eco-Nano-Technologies and advanced materials, Bioeconomy and Energy, environment and climate change smart specialization domains, however the highest percentage of contracted values was secured by national institutes of research and development type beneficiaries.

Action 1.2.3 - Knowledge transfer partnerships

- Section G (Knowledge transfer partnerships) 54 projects were funded, with a total value of 129.87 million euro (out of which 90.64 million euro in ERDF).

The most salient beneficiaries in this action were the national institutes of research and development (54%) and universities (44%) implementing projects under IT&C, space and security, followed by the Eco-Nano-Technologies and advanced materials smart specialization domains; however, the highest percentage of contracted values was secured by microenterprises.

Overall, we can conclude that the most salient beneficiaries per smart specialization domains and the national priority domain, are as follows:

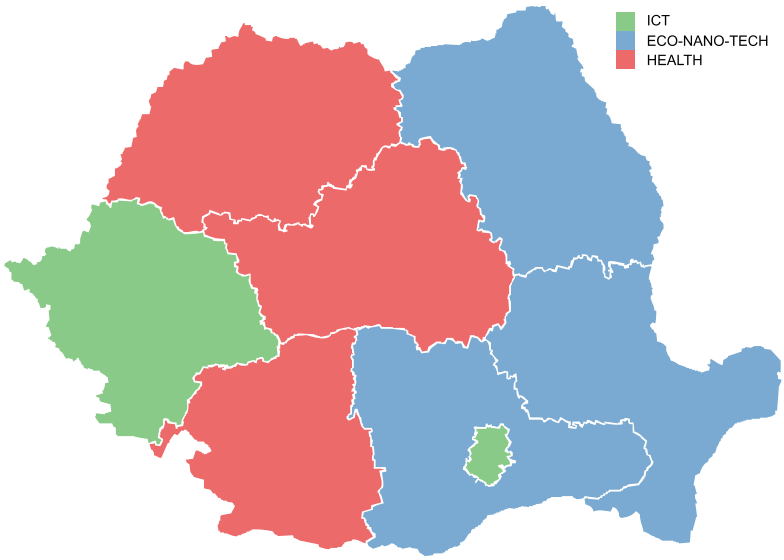
- Bioeconomy (Associations)
- IT&C, space and security (Microenterprise, Small Enterprises, Foundations, NGOs)
- Energy, environment and climate change (Medium Enterprises, Another type of beneficiaries: public institutions with research activities; NGOs)
- Eco-Nano-Technologies and advanced materials (National Institutes for Research and Development, Universities, Large Enterprises, Research Institutes of the Romanian Academy)
- Health (hospitals)

Objective 2: Regional competencies

One of the goals of Priority Axis 1 (PA1) in the Competitiveness Operational Programme is to support the smart specialization priorities, taking into account the research orientation towards the economic needs and correlating them to the founding instruments active in the national R&D plan. The analysis of the funded projects identifies the existence of research critical mass, resources and competencies in the eight development regions in various smart specialization areas, as seen in Figure 1.

Smart specialization funded projects by development regions

Figure 1



For the purpose of economic growth and development, as well as for policy reasons, the analysis provides insight on regional competencies in terms of smart specialization domains: Bucharest-Ilfov and the West region have competencies in implementing projects under the IT&C, space and security area, the South-West, Centre and North-West have competencies in implementing projects under the Health area, whereas the rest of the regions developed competencies in implementing projects funded under the Eco-Nano-Technologies and advanced materials. The other two specialization domains, Bioeconomy and Energy, environment and climate change did not stand out as primary competencies in any area.

Out of the 331 projects funded under the PA1 of the Competitiveness Operational Programme 2014-2020, 146 projects (44% of all projects) are connected to the Classification of Activities in the National Economy (NACE code). The tendencies we have identified are as follows:

Under the IT&C, space and security domain, there is a preference for Service activities in IT (first two digits of NACE code 62), accounting for 14.38% of all national projects. Under the Health domain, there is a preference for Human health activities (first two digits of NACE code 86) accounting for 15.7% of all national projects.

Bucharest-Ilfov region, as the location of project implementation is the host of 10.96% of the projects classified under the Service activities in IT NACE.

Objective 3: Non eligible value of the project by action type, the type of projects and legal status.

In the Competitiveness Operational Programme, we refer to the non-reimbursable financial aid (NFA) as the contribution of the funding body to finalize the investment and can cover partially or fully the eligible value of a project. Investigating the non-eligible value of the funded projects under the PA1 of the COP, the weight of the non-eligible depends on the action under which the project is implemented. Therefore, we emphasize for each action, the type of beneficiary that registered the highest percentage value of the non-reimbursable financial aid and the kind of projects in which it was implemented.

One can note the following general characteristics for all types of beneficiaries and for all types of actions financed through POC 2014 - 2020: the obligation of ineligible expenses with the project audit and erroneous classification of ineligible expenses on eligible activities.

Under the Action 1.1.1 - Large R&D infrastructures - 8 projects were funded for public R&D institutions/universities; the percentage of the non-eligible value was 5.28% of the NFA (87.79 million euro) for universities, for type F projects. Under this action the activities identified as sources that led to the increase of non-eligible expenses are: the complexity and necessity of all the actions for the implementation of the project, projects that had a total value above the maximum eligible value (allowed by the applicant's guide), the difference being considered an ineligible expense supported by the beneficiary or the incorrect classification of certain categories such as approvals and agreements for ineligible expenses.

Under the Action 1.1.2 - Development of networks for R&D centres - 10 projects were funded for universities; the percentage of the non-eligible value was 3.35% of the NFA (56.15 million euro). The highest percent of non-eligible value was recorded for Cloud type projects. Under this action the activities that led to the increase in the non-eligible expenses are: non-compliance with the thresholds established by the applicant's guide, of 15% for indirect costs and 10% management costs, respectively; the total value of the project was above the maximum eligible value (allowed by the applicant's guide) the difference being considered as ineligible expenditure, supported by the beneficiary and also the VAT that was classified by the applicant as ineligible.

Under the Action 1.1.3 - Synergies with Horizon 2020 - 9 projects were funded with universities as beneficiaries; the percentage of the non-eligible value was 0.17% of the NFA (2.03 million euro). The highest percent of non-eligible value was recorded for the RO-ECSEL type projects. Under this action, the beneficiary of type enterprise had the obligation to introduce the following ineligible costs in the project: expenses for information and publicity regarding the project and expenses for project management. If the beneficiary of a project under this action is a research entity (institute or university) and it is part of a consortium, then the publicity expenses of the project and management expenses become ineligible for the research entity as well.

Under the Action 1.1.4 - Attract researchers from abroad - 27 projects were funded with universities as beneficiaries; the percentage of the non-eligible value was 4.48% of the NFA (8.68 million euro). The highest percent of non-eligible value was recorded for type E projects. Under this action the ineligible expenses have the same reason as described under the previous action and it was also observed that failure to comply with the thresholds set by the applicant's guide was identified.

Under the Action 1.2.1 - Stimulate the demand of enterprises for innovation - 10 projects were funded with medium enterprises as beneficiaries; the percentage of the non-eligible value was 20.55% of the NFA (4.39 million euro). The highest percent of ineligible cost value was recorded for the Innovative technological type (PTI) projects. The ineligible cost values under this action are due to the VAT, management or publicity expenses that become ineligible or the fact that public procurement of consulting services in the field of innovation exceed the value maximum of 200,000 euro, established as aid for the beneficiary.

Under the Action 1.2.3 - Knowledge transfer partnerships - 29 projects were funded with national research institutes as beneficiaries; the percentage of the non-eligible value was 1.48% of the NFA (0.87 million euro). The ineligible costs under this action are mostly due to the obligation of the beneficiary to support the ineligible expenses of the project audit.

Objective 4: Eligible salaries and non-reimbursement salaries tendencies.

We have investigated the difference between median salaries -- both eligible and non-reimbursement -- in terms of action type, smart specialization domains, legal status, applicant headquarters, and the location of project implementation. To this end, we have conducted Kruskal-Wallis tests followed by pairwise Mann-Whitney U-test tests with Bonferroni correction available in Appendix A.

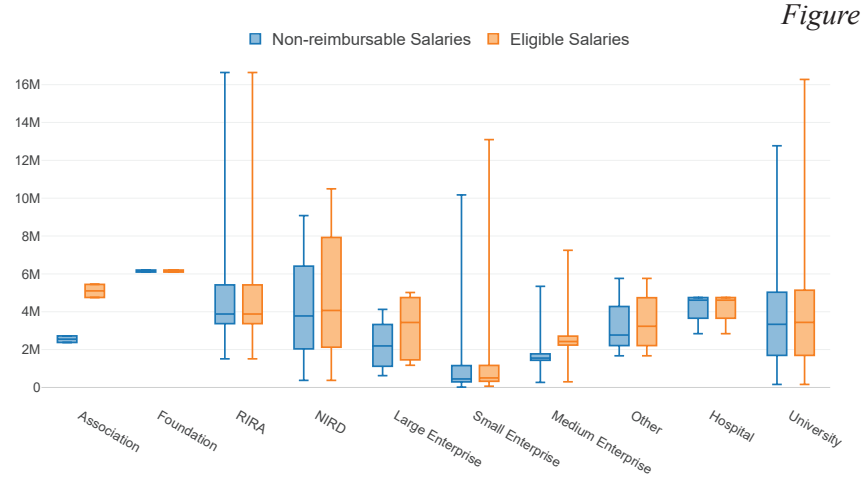
We found no significant difference in terms of applicant headquarters, location of project implementation, and smart specialization domains.

There are significant differences between eligible salaries in terms of legal status as shown in Figure 2. The eligible salaries applied within microenterprises are significantly lower than those applied within medium enterprises, universities, the National Institute for Research and Development, and the Research Institute of the Romanian Academy. This finding is expected, as the max eligible monthly salary stipulated in the COP guidelines is based on education level and professional rank (ex: research assistant, PhD students, tenure professor, manager/director etc), thus universities and other R&D institutes will have a higher median salary requirement.

Similarly, the non-reimbursement salaries are significantly different between the legal status categories. The non-reimbursement salaries applied within microenterprises are significantly lower than those applied within universities, the National Institute for Research and Development, and the Research Institute of the Romanian Academy. However, there is no significant difference between the non-reimbursement salaries applied within microenterprises and medium enterprises.

Eligible or non-reimbursable salaries by legal status

Figure 2



There are significant differences between eligible salaries in terms of action type as shown in Figure 3.

Eligible or non-reimbursable salaries by action type

Figure 3

The eligible salaries significantly differ between most action types with the eligible salaries applied for "*Knowledge transfer partnerships"* being higher than those applied for all the other action types. Similarly, the non-reimbursement salaries applied for "*Knowledge transfer partnerships"* are significantly higher than those applied for all the other action types.

3. RESULTS

This study sought to investigate funding granted through the Competitiveness Operational Programme 2014-2020, which represents the national implementation of Romania's Partnership Agreement. We have conducted our analysis in terms of four evaluation objectives. The main conclusions are as follows:

Objective 1: To identify the common tendencies among funds recipients

The universities and the national institutes of research and development were the most active in securing funds among all action types. Out of the six action types implemented under COP, universities were over-represented in four categories: Action 1.1.2 - Development of networks for R&D centres, Action 1.1.3 - Synergies with Horizon 2020, Action 1.1.4 - Attract researchers from abroad, Action 1.2.3 - Knowledge transfer partnerships.

The national institutes of research and development are most active among Action 1.1.1 - Large R&D infrastructures, while Action 1.2.1 -

Stimulate the demand of enterprises for innovation is approached mainly by microenterprises.

Objective 2: To identify the regional competencies

The competencies are heterogeneously distributed among NUTS regions.

- Bucharest-Ilfov and the West region -- competencies in implementing projects under the IT&C, space and security area
- the South-West, Centre and North -West regions -- competencies in implementing projects under the Health area
- the North East, South East, and North-West regions -- competencies in implementing projects funded under the Eco-Nano-Technologies and advanced materials.

Objective 3: A description of the non-eligible value of the projects.

Since the financing entity did not foresee all eligible expenses and the beneficiaries still needed to carry out a series of activities necessary for the efficient implementation of the project, we identified non-eligible salaries for all action types. The ineligible expenses found in some projects included additional amounts for the achievement of the objectives established by the project.

Objective 4: A description of eligible salaries and non-reimbursement salaries tendencies.

There are significant median differences between the eligible salaries recorded by microenterprises and the eligible salaries incurred by the medium enterprises, universities, the National Institute for Research and Development, and the Research Institute of the Romanian Academy. The eligible and non-reimbursement salaries applied for "Knowledge transfer partnerships" are significantly higher than those applied for all the other action types. No significant difference between eligible or non-reimbursement salaries in terms of applicant headquarters, or the location of project implementation, or smart specialization domains was identified. However, eligible salaries tend to increase as the total value of the project increases.

4. DISCUSSION

Existing EU policies seek to support national research, development, and innovation activities and, through the structural funds, operational programs aim to promote competitiveness and innovation (EUROPEAN COMMISSION, 2010).

The Competitiveness Operational Programme is anchored in the general strategic objective of NSRDI to increase *the competitiveness of the Romanian economy through innovation.* To this end, it endeavours to support the performance of economic operators on overall value chains. The national strategy supports the conversion from cost-based to innovation-based competitiveness. This transition means to grow the ability of companies to attract state-of-the-art technology, to adjust these technologies to societal needs, and to develop technologies or services enhancing added value. The specific objectives of COP, strongly anchored in the Romanian RDI strategic document and responding to the strategic and cross-cutting objectives of NSRDI, are as follows:

- Increasing scientific capacity in the fields of smart specialization and health;
- Increasing involvement in research at the EU level;
- Increasing private investment in RDI;
- Increasing the transfer of knowledge, technology, and staff with RDI skills among the public environment research and the private one (Competitiveness Operational Programme, 2014).

The contracted projects tackled all domains and subdomains of smart specialization identified in the national strategic framework of RDI as a priority domain of national interest (bioeconomy; ICT, space and security; energy, environment and climate change, eco-nano-materials, health), increasing involvement in research at EU level. COP has supported the consolidation of investments in RDI actions within enterprises, including both the financing of large infrastructures and the implementation of collaboration activities within large innovative clusters, as well as the collaboration for knowledge and technology transfer in smaller partnerships (research organization - enterprise), having the objective of passing the research results to the business environment, able to increase the competitiveness on the market and to bring added value in the global value chain. The RDI component within the COP increases the mobilization of public and private actors in reaching the European objectives taking into account the key role of innovation and research in the transition to a knowledge-based, competitive and sustainable economy by supporting areas of smart specialization. Smart Specialisation plays a major role in stimulating RDI at the national and regional levels in both less and more developed regions.

The contribution of universities in securing funds among all action types aligns with European standards. Under the smart specialization concept, the EU Cohesion Policy underlines the role of universities in formulating innovation strategies and identifying regional priorities (McCann and Ortega-Argilés, 2015; Muller et al., 2017). It has been mentioned that Research and

Innovation Strategy for Smart Specializations (RIS3) strategies, an ex-ante prerequisite to access the European Regional Development Funds (ERDF), can facilitate aligning universities' research with regional needs (Charles et al., 2014; Fonseca and Salomaa, 2020).

Salaries are similar among development regions, action types, and institutions of comparable dimensions. They only tend to increase as the total value of the project increases. Financial incentives are determinant for the number and the quality of human resources in research, given that new jobs have been created and existing ones have been maintained. This contributes to the cross-cutting objectives of NSRDI (SO5) to incentivize a critical mass of researchers that will transform research and development and innovation into a factor of economic growth, guaranteeing fast and sustainable evolution, from the quantitative and qualitative perspective of the human resource and the innovation and development results.

Our evaluation reveals a good fit between the above-mentioned specific objectives of COP and the actual implementation results. A flaw of COP implementation arises from the deficient anticipation of eligible expenses.

5. **References**

1. Autoritatea de Management pentru Programul Operațional Competitivitate (n.d.). Available at: https://mfe.gov.ro/programe/autoritati-de-management/am-poc/ (accessed 27 October 2021).
2. **Bachtler J and Michie R** (1995) *A New Era in EU Regional Policy Evaluation? The Appraisal of the Structural Funds*. Regional Studies 29(8). Routledge: 745–751. DOI: 10/cwqcrv.
3. **Charles D, Kitagawa F and Uyarra E** (2014) *Universities in crisis?—new challenges and strategies in two English city-regions*. Cambridge Journal of Regions, Economy and Society 7(2): 327–348. DOI: 10/gm8m2z.
4. Competitiveness Operational Programme (2014). Available at: https://ec.europa.eu/regional_policy/en/atlas/programmes/2014-2020/romania/2014ro16rfop001 (accessed 10 November 2021).
5. EUROPEAN COMMISSION (2010) *Europe 2020: A European strategy for smart, sustainable and inclusive growth*. Available at: https://ec.europa.eu/eu2020/pdf/COMPLET%20EN%20BARROSO%20%20%20007%20-%20Europe%202020%20-%20EN%20version.pdf (accessed 27 October 2021).
6. **Fonseca L and Salomaa M** (2020) *Entrepreneurial Universities and Regional Innovation: Matching Smart Specialisation Strategies to Regional Needs?* IGI Global. DOI: 10.4018/978-1-7998-0174-0.ch014.
7. **McCann P and Ortega-Argilés R** (2015) *Smart Specialization, Regional Growth and Applications to European Union Cohesion Policy*. Regional Studies 49(8). Routledge: 1291–1302. DOI: 10/f3mg88.
8. **Muller E, Zenker A, Hufnagl M, et al.** (2017) *Smart specialisation strategies and cross-border integration of regional innovation systems: Policy dynamics and challenges for the Upper Rhine*. Environment and Planning C: Politics and Space 35(4). SAGE Publications Ltd STM: 684–702. DOI: 10/gm8m2v.
9. National Strategy for Research, Development and Innovation 2014-2020 (2018). Available at: https://ec.europa.eu/growth/tools-databases/regional-innovation-monitor/policy-document/macroregiunea-doi/national-strategy-research-development-and-innovation-2014-2020 (accessed 27 October 2021).

Appendix A

Objective 4 - Statistical Analysis

We tested the statistical significance of median differences through non-parametric models. Considering the increasing emphasis on effect sizes, results are reported both on means of p-values and effect sizes. To this end, we have conducted Kruskal-Wallis tests followed by pairwise Mann-Whitney U-test tests with Bonferroni correction. We conducted our data analysis using R statistical software, version 4.1.1.

Applicant headquarters

There is a significant difference between the eligible salaries in terms of applicant headquarters (Kruskal-Wallis H = 17.991, df = 7, p-value = 0.01201, eta squared = 0.0340). However, pairwise tests with Bonferroni correction reveal no significant difference.

There is no significant difference between non-reimbursement in terms of applicant headquarters (Kruskal-Wallis H = 16.817, df = 7, p-value = 0.01862, eta squared = 0.0304). However, pairwise tests with Bonferroni correction reveal no significant difference.

Project implementation

There is no significant difference between the eligible salaries in terms of project implementation (Kruskal-Wallis H = 13.265, df = 7, p-value = 0.06592, eta squared = 0.0194).

There is no significant difference between non-reimbursement in terms of project implementation (Kruskal-Wallis H = 11.971, df = 7, p-value = 0.1015, eta squared = 0.0154).

Smart specialization

The results reveal a significant difference between the eligible salaries applied within the five specialization domains (Kruskal-Wallis H = 14.319, df = 4, p-value = 0.006343, eta squared = 0.0317) with further investigation underling a significant difference between the eligible salaries applied for *Bioeconomy* and *Health* (p-value = 0.032). However, these results are associated with a small effect size. Similarly, we have identified a significant difference between the non-eligible salaries applied within the five specialization domains. The test revealed a significant difference between the five specialization domains (Kruskal-Wallis H = 11.936, df = 4, p-value = 0.01783, eta squared = 0.0243). However, pairwise tests with Bonferroni correction reveal no significant difference. Considering these and the small effect sizes, we can conclude that there is no compelling difference between

the salaries (eligible or non-reimbursement) salaries applied within the five specialization domains.

Legal status

There are significant differences between eligible salaries in terms of legal status (Kruskal Wallis H = 119.39, df = 9, p-value < 2.2e-16, eta squared = 0.377). The eligible salaries applied within microenterprises are significantly lower than those applied within medium enterprises (p-value=0.03), universities (p-value=6.7e-15), the National Institute for Research and Development (p-value=7.2e-15), and the Research Institute of the Romanian Academy (p-value<0.01).

The non-reimbursement salaries are significantly different between the legal status categories (Kruskal Wallis H = 124.46, df = 9, p-value < 2.2e-16, eta squared = 0.394). The non-reimbursement salaries applied within microenterprises are significantly lower than those applied within universities (p-value=8.1e-16), the National Institute for Research and Development (p-value=3.9e-15), and the Research Institute of the Romanian Academy (p-value<0.01).

Action type

There are significant differences between eligible salaries in terms of action type (Kruskal Wallis H = 171.71, df = 5, p-value < 2.2e-16, eta squared = 0.561). "*Knowledge transfer partnerships"* eligible salaries are higher than those applied for all the other action types (Large R&D infrastructures: p-value = 0.00030; Development of networks for R&D centres: p-value = 8.6e-06; Synergies with Horizon 2020: p-value = 7.9e-09; Attract researchers from abroad: p-value = 7.7e-05; Stimulate the demand of enterprises for innovation: p-value < 2e-16).

Non-reimbursement salaries are significantly different between action types (Kruskal-Wallis H = 11.936, df = 4, p-value = 0.01783, eta squared = 0.0243). The non-reimbursement salaries applied for "*Knowledge transfer partnerships"* are significantly higher than those applied for all the other action types (Large R&D infrastructures: p-value = 0.00057; Development of networks for R&D centres: p-value = 8.6e-06; Synergies with Horizon 2020: p-value = 7.9e-09; Attract researchers from abroad: p-value = 0.00832; Stimulate the demand of enterprises for innovation: p-value < 2e-16).

Correlation of Final Consumption of Energy Used, Renewable and Biofuels, with Gross Domestic Product

Assoc. prof. Nicolae MIHĂILESCU, PhD. (n.mihailescu@yahoo.com)
„Hyperion" University of Bucharest

Professor Florinel-Marian SGĂRDEA, PhD. (sgardeafm@gmail.com)
The Bucharest Academy of Economic Studies (ASE)

Assoc. prof. Claudia CĂPĂȚÎNĂ, PhD. (claudiacapatana@yahoo.com)
„Hyperion" University of Bucharest

ABSTRACT

Energy is a current topic of priority importance for human activity. The political and economic decisions implemented by each state are rigorously based on complex analyzes to assess the level of pollution, the ability to respond to pollution limitation and the greenhouse effect. The development of human society, the socio-economic progress and the standard of living depend, first of all, on the ever-increasing consumption of energy.

This article presents an econometric model of the final consumption of energy used, renewable and biofuels in the 27 countries of the European Union depending on the size of the gross domestic product. The economic power of each state identified by the value of gross domestic product determines a sustainable response to finance the production and distribution of energy from renewable sources and biofuels.

The study was conducted on the basis of information provided by Eurostat and the International Monetary Fund for 2019.

Periodic review of the research on an updated statistical data system is a solution to be considered as a way of operative knowledge of the achievements related to the analyzed field.

The research presented applies a methodology that is rigorously grounded in statistics, econometrics and macroeconomics and can also be used as a support for applied information to develop econometric models with viability and practical utility for decision makers.

EViews software was used to define the econometric model.

The study concludes with conclusions on the viability of the econometric model of the correlation of the final consumption of renewable energy and biofuels in the European Union-27 with gross domestic product, with an indisputable utility for substantiating government decisions aimed at economic policy of monitoring and limiting pollutants.

The general trend at European level is the annual increase in the final consumption of renewable energy and biofuels as a result of the concerns and measures applied in the countries of the Union, both political and especially economic, organizational, scientific and technological.

Keywords*: renewable energy and biofuels, econometric model.*
JEL classification*: C13*

INTRODUCTION

The final consumption of energy used, renewable and biofuels is directly dependent on the economic potential of the state, the decision of the administrative and political leadership to regulate the energy field, production and distribution.

It is a generally accepted economic logic that the „Gross Domestic Product" is considered as the indicator that synthetically expresses the economic potential of a state and is, at the same time, the indicator that summarizes economic growth when its evolution is characterized by positive rhythms. By increasing the total gross domestic product and per inhabitant, the material and financial basis necessary for the fulfillment of the objectives of existence and sustainable development of all fields of economic and social life is ensured.

In order to achieve this major goal of functioning of the national economy, it is necessary to carry out studies with a viable methodological structure, with a correctly formulated economic motivation, to provide the necessary information support to identify a useful and real diagnosis of the performance of "green energy" production and consumption and the substantiation, on this basis, of some economic, budgetary, fiscal and legislative policy decisions, decisions that target both immediate time segments but also longer periods of time.

The motivational considerations presented can provide the opportunity to develop an econometric model that expresses in a synthetic (mathematical) way the correlation of the final consumption of energy used, renewable and biofuels with the gross domestic product by applying a rigorous econometric modeling methodology.

REFERENCE LITERATURE

The analysis of the dynamics of the final consumption of renewable energy and biofuels in the states of the European Union - 27 joins the numerous econometric modeling works that have been presented in articles and specialized papers in the country and abroad.

All the studies to which we refer use a modeling methodology, rigorously based on economic theory, mathematical statistics, probability theory and statistical inference, respectively. Particular cases of analysis are treated separately by mathematical modeling of the dynamics of economic variables as well as the formation of interdependencies between variables formed in a systemic format and also provide appropriate solutions for model validation.

In this sense, the most recent works are relevant, published by Anghelache, C., Anghel, M.G., Manole, A. (2015) - "Economic, financial-banking and informatics modeling", Artifex Publishing House, Bucharest[1]; 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[2]; Andrei, T., Bourbonais, R. (2008) - „Econometrics", Economic Publishing House, Bucharest[3].

In the paper elaborated by Mihăilescu, N. (2014) - "Statistics and Statistical Bases of Econometrics", Transversal Publishing House, Bucharest[4], unifactorial and multifactorial models are defined and validated, both at macroeconomic and microeconomic level, both in Romania as well as the states of the European Union.

A similar topic that refers to the complex analysis of the evolution of the money supply with the help of interdependent models is treated by Mihăilescu, N. (2019) - „Analysis of economic-financial activity - Research methodologies, case studies solved to substantiate economic-financial decisions and tests of knowledge", Transversal Publishing House, Bucharest[5].

Other reference works are: Mihăilescu, N., Căpăţână, C. (2018) - „The reversible impact of the dynamics of the gross domestic product with the imports and exports of goods and services of Romania, Romanian Statistical Review - Supplement/ Nr. 11 and No. 12/2018[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, Publishing House Artifex, Bucharest, Vol. 2 / No.4 / 2011[7].

The mentioned works present in the context of the scientifically based methodology of econometrics, statistical legitimacies expressed by regression or trend equations that are formed to express the reality of dynamic economic processes, on quarterly and annual time segments or as interdependence between economic variables, in profile static - at a given time or representative time segment, both macroeconomically and microeconomically level. Models of demographic or social variables depending on the size of an economic variable are also presented.

RESEARCH METHODOLOGY AND STUDY DATABASE

The research methodology of the correlation of the final consumption of energy used, renewable and biofuels, from 2019 with the gross domestic product for 27 states that make up the European Union is based on econometrics and is based on the following steps:

- Graph the correlation of the indicators in the database and choose the mathematical form of the model.

- The coefficients of the econometric model are estimated by applying the least squares method and their statistical significance is verified using „*Criterion t*",

- The viability of the model is assessed by specific tests using the following criteria: „*Criterion F*", „*Jarque-Bera Criterion*", „*Durbin-Watson Criterion*" and „*White Test*".

- It also quantifies the „power" of the model for calculating predictable levels of final consumption of energy used, renewable and biofuels using „*Theil's Irregularity/Inequality Coefficient*" as well as the relative expression „*Standard error estimation of the equation of regression*",

- Forecast levels are estimated, as a point value and as a guaranteed confidence interval with a probability of at least 95%.

The methodology used to develop and certify the viability of the models is applied using the Eviews software.

The database that is used to carry out the proposed study is provided by Eurostat and the International Monetary Fund with reference to the year 2019 and respectively to the 27 states that make up the European Union, Table 1.

List of European Union countries - 27 after final consumption of energy used, renewable and biofuels, total and per capita, total GDP and per capita and Population (2019)

Table 1

No.	Member State	Final consumption of energy used, renewable and biofuels (thousand tons of oil equivalent)	GDP (nominal), (million US dollars)[1]	The population	Final consumption of energy used, renewable and biofuels per capita (tons of oil equivalent)	Gross domestic product per capita ($)
1	Germany	16,618.056	4,000,386	84,052,061	0.197711	47,594.15
2	France	14,836.710	2,775,252	65,418,030	0.226798	42,423.35
3	Italy	10,912.230	2,072,201	60,776,531	0.179547	34,095.41

4	Spain	7,202.379	1,425,865	46,770,000	0.153996	30,486.74
5	Poland	6,418.452	586,015	38,511,824	0.166662	15,216.50
6	Romania	3,831.610	239,851	20,121,641	0.190422	11,920.05
7	Lower Countries	1,988.634	912,899	16,787,689	0.118458	54,379.08
8	Greece	1,701.095	219,097	11,350,118	0.149875	19,303.50
9	Belgium	1,948.804	533,153	11,071,483	0.176020	48,155.52
10	Portugal	2,904.002	238,510	10,562,178	0.274943	22,581.52
11	Czech Republic	3,242.387	242,052	10,436,560	0.310676	23,192.70
12	Hungary	1,846.200	155,703	9,906,000	0.186372	15,718.05
13	Sweden	8,210.396	551,135	9,573,466	0.857620	57,569.01
14	Austria	4,110.477	457,637	8,602,112	0.477845	53,200.54
15	Bulgaria	1,424.833	64,963	7,351,234	0.193822	**8,837.020**
16	Denmark	1,675.684	350,874	5,580,516	0.300274	62,874.83
17	Finland	6,763.272	275,321	5,421,827	**1.247416**	50,780.12
18	Slovakia	1,236.644	106,585	5,414,937	0.228376	19,683.52
19	Ireland	487.697	424,635	4,588,252	0.106293	92,548.32
20	Croatia	1,151.904	60,688	4,290,612	0.268471	14,144.37
21	Lithuania	723.309	53,323	3,043,429	0.237663	17,520.70
22	Latvia	1,037.812	34,881	2,217,000	0.468115	15,733.42
23	Slovenia	641.394	54,242	2,062,455	0.310986	26,299.73
24	Estonia	442.860	30,312	1,294,455	0.342121	23,416.80
25	Cyprus	197.276	24,492	838,916	0.235156	29,194.82
26	Luxembourg	161.030	68,770	537,853	0.299394	**127,860.2**
27	Malta	31.875	14,505	452,515	**0.070440**	32,054.19

Source: Eurostat,[1] Data provided by the International Monetary Fund (2019)

The final consumption of energy used, renewable and biofuels of 27 component states of the EU, in 2019, clearly highlights the existence of a state of pronounced heterogeneity of European states assessed on the basis of the size of the coefficient of variation. Finland has the highest per capita consumption, 1.247416 tons of oil equivalent and Malta the lowest level 0.070440. This state of affairs is described in detail and edifying by appropriate statistical indicators systematized in Table 2 for total values (column 1) and for values per capita (column 2 - solution adopted to ensure the comparability of consumption size). It is clear that the information on this energy consumption raises the question of meeting the challenge of greater involvement of European states in coordinating policies to support the rapprochement of achieving the desired result, reducing environmental pollution.

The situation of the 27 states that make up the European Union in terms of the degree of homogeneity / heterogeneity in terms of the size of the gross domestic product per capita, Table 2 (column 3) is also attested with a pronounced heterogeneity based on the coefficient of variation which is 72.4579%, much higher compared to a sufficiently restrictive limit of 10%.

Statistical description of the set of values for the final consumption of energy used, renewable and biofuels, total și per capita și GDP per capita of 27 EU states. in 2019

Table 2

Statistical indicators	Final consumption of energy used, renewable and biofuels (thousand tons of oil equivalent)	Final consumption of energy used, renewable and biofuels per capita (tons of oil equivalent)	GDP per capita (nominal), ($)
0	1	2	3
Mean	3,768.408	0.295388	36,917.93
Median	Hungary: 1,846.200	Slovakia: 0.228376	Cyprus: 29,194.82
Maximum	Finlanda: 16,618.06	Finlanda: 1.247416	Luxemburg: 127,860.2
Minimum	Malta: 31.87500	Malta: 0.070440	Bulgaria: 8,837.020
Std. Dev.	4,441.705	0.244274	26,749.96
Coefficient of variation (%) = (Std. Dev/Mean) 100	117.87%	82.696%	72.4579%
Skewness	1.625991	2.719834	1.758785
Kurtosis	4.794893	10.43553	6.334543
Jarque-Bera	15.52166	95.48672	26.42903
Probability (*J-B*)	0.000426	0.000000	0.000002
Sum	101,747.0	7.975471	996,784.2
Sum Sq. Dev.	(5.13E+08)	1.551411	(1.86E+10)
Observations	27	27	27

As mentioned, the statistical series of values on the final consumption of energy used, renewable and biofuels, per inhabitant, shows a pronounced asymmetric arrangement. Table 3 identifies a group of 12 states that report a final consumption of energy used, renewable and biofuels, per capita up to 0.2 tons of oil equivalent and another 11 states have a consumption of between 0.2 and 0.4. These two groups represent 85.19% of the total states.

The states with a final consumption of energy used, renewable and biofuels, per capita, exceeding 0.4 tons of oil equivalent are 4, Finland (1.247416), Sweden (0.857620), Austria (0.477845) and Latvia (0.468115), with an overall proportion of 14.81%.

The graphical representation in Figure 1 illustrates the situation in Table 3 with the obvious asymmetry that characterizes the disposition of the 27 states of the European Union after the final consumption of energy used, renewable and biofuels, per inhabitant.

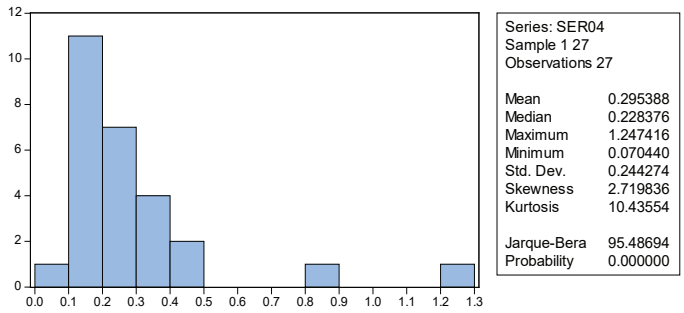
Grouping the statistical series of values regarding the final consumption of energy used, renewable and biofuels, per capita, of 27 EU states in 2019 (tons of oil equivalent)

Table 3

Sample: 1 - 27: Number of observations: n = 27: Number of groups: 5				
Value (tons of oil equivalent)	Count	Percent	Cumulative Count	Cumulative Percent
[0.0 – 0.2)	12	44.44	12	44.44
[0.2 – 0.4)	11	40.74	23	85.19
[0.4 – 0.6)	2	7.41	25	92.59
[0.8 – 1.0)	1	3.70	26	96.30
[1.2 – 1.4)	1	3.70	27	100.00
Total	27	100.00	27	100.00

Graphical representation of the statistical series of values on final consumption of energy used, renewable and biofuels, per capita, of 27 EU states in 2019, (tons of oil equivalent)

Figure 1

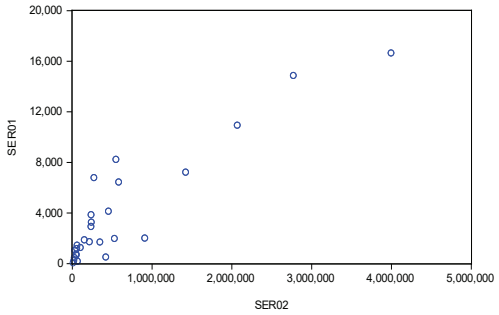


Defining the econometric model of the correlation of the final consumption of energy used, renewable and biofuels, with the gross domestic product, calculating the indicators of econometric representation, testing their statistical significance and comments

The analysis of the graphical representation of the correlation between the two variables (Figure 2) shows that as the gross domestic product increases, so does the final consumption of energy used, renewable and biofuels, which confirms a direct link, and the points tend to grouped around a line, which justifies the option for a linear function. Under these conditions, the regression equation $\hat{y} = a + bx$ will represent the estimated levels of the endogenous variable (y) as a function of the exogenous variable (x) and the equation that exposes synthetically to the real levels is: $y = a + bx + u$ where u is the residual variable.

Correlogram of final consumption of energy used, renewable and biofuels expressed in thousand tons of oil equivalent (SER01) by gross domestic product - million US dollars (SER02), for 27 countries of the European Union, in 2019

Figure 2



The indicators that ensure an analytical and at the same time complex characterization of the econometric model (estimated regression equation) are presented in Table 4.

Synoptic table of the system of econometric representation indicators for the model of the correlation of the final consumption of energy used, renewable and biofuels with the gross domestic product

Table 4

Dependent Variable: ***y* = Final consumption of energy used, renewable and biofuels** (thousand tons of oil equivalent)				
Method: Least Squares				
Sample: 1-27; Included observations: *n* = 27				
The regression equation: y = a + b · x + u; y = 1,232.346 + 0.004287 · x + u				
Variable	Coefficient	Std. Error	*t*-Statistic	Prob.
x = GDP - nominal (million US dollars) „b”	0.004287	0.000386	11.11208	0.0000
C = model constant „a”	1,232.346	424.3097	2.904356	0.0076
R-squared	**0.831625**	Mean dependent var: $\bar{y}$		3,768.408
Adjusted *R*-squared	0.824890	S.D. dependent var		4,441.705
S.E. of regression	**1,858.682**	Jarque-Bera		4.634458
Sum squared resid	86,367,436	Prob. (*J-B*)		**0.098546**
Correlation ratio: R=$\sqrt{R^2}$	**0,911935**	Hannan-Quinn criter.		17.99285
F-statistic	123.4782	Durbin-Watson stat		1.848242
Prob (*F*-statistic)	0.000000	*Theil Inequality Coefficient*		**15.9134 %**
Heteroskedasticity Test: White				
F-statistic = 0.680581;		Prob. *F* (2, 24)		0.5158
$n \cdot R^2 = \chi^2 -$ statistic $=$ 1.449121;		Prob. Chi-Square (2)		0.4845

Figure 3 provides visual information on how the three components defining the model (actual and estimated endogenous variable data and residues, respectively) are located for each of the 27 states in 2019. The graphical form confirms the status of the results in Table 5.

It can also be estimated that the size of the residue does not exceed the estimate of the limit error ($\hat{\Delta}$), resulting from the product of the critical value of $t\text{-}_{table} = \pm 2.060$, for a probability of 95% (significance threshold of 5% is arranged bilaterally) and 25 degrees of freedom (based on the Law of Student Distribution), $f = n - k = 27 - 2 = 25$, with the estimation of the standard error of the regression equation, $\hat{\sigma}_{y,\hat{y}} = \pm$ 1,858.682, graphically presented situation in the last column of Table 5, except Sweden and Finland. These statistical findings support with sufficient confidence the viability of the model of correct representation of reality.

($\hat{\Delta} = 2.060 \cdot 1{,}858.682 = \pm$ 3,828.885 thousand tons of oil equivalent)

Graphical presentation of residues (Residual), real levels - calculation basis - (Current) and estimated levels (Fitted) based on the linear regression equation of the correlation of the final consumption of energy used, renewable and biofuels, with the gross domestic product

Figure 3

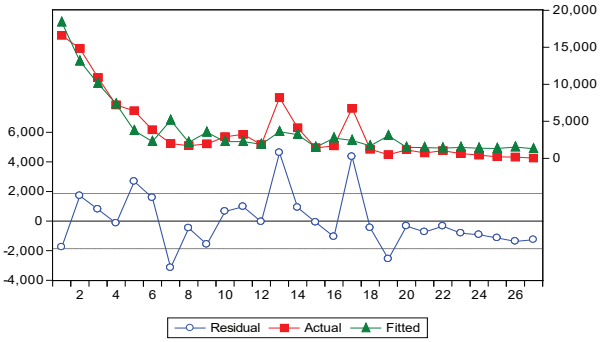


Table 5 shows the actual levels of the dynamics of final consumption of energy used, renewable and biofuels, the estimated levels based on the simple linear regression equation, and the series of error term levels. The range of residues in the last column of the table gives the image of a corresponding alternative arrangement of the error term compared to the zero size. This confirms, in graphical form, the existence of the state of non-autocorrelation of the values of the error term, identified as a numerical dimension by the size of the Durbin-Watson statistical coefficient (DW = 1.848242) and consequently it is estimated that the model is correctly developed.

Series of real levels (calculation basis), of the levels estimated on the basis of the regression equation, regarding the correlation of the final consumption of energy used, renewable and biofuels, with the gross domestic product and the residual term range - (Linear unifactorial econometric model) (thousand tons of oil equivalent)

Table 5

No.	Member State	Actual (thousand tons of oil equivalent) y	Fitted (thousand tons of oil equivalent) $\hat{y}$	Residual $u = y - \hat{y}$	Residual plot $\pm\hat{\sigma}_{y,\hat{y}} = \pm 1{,}858.682$ $-\hat{\sigma}_{y,\hat{y}}$ $+\hat{\sigma}_{y,\hat{y}}$
1	Germany	16.618,1	18.381,0	-1.762,93	* | .
2	France	14.836,7	13.129,1	1.707,57	. | *.
3	Italy	10.912,2	10.115,3	796,886	. | * .
4	Spain	7.202,38	7.344,67	-142,287	. * .
5	Poland	6.418,45	3.744,44	2.674,01	. | . *
6	Romania	3.831,61	2.260,53	1.571,08	. | *.
7	Lower Countries	1.988,63	5.145,71	-3.157,08	* . | .
8	Greece	1.701,10	2.171,56	-470,464	. * | .
9	Belgium	1.948,80	3.517,84	-1.569,03	.* | .
10	Portuga	2.904,00	2.254,78	649,224	. | * .
11	Czech Republic	3.242,39	2.269,96	972,425	. | * .
12	Hungary	1.846,20	1.899,81	-53,6055	. * .
13	Sweden	8.210,40	3.594,92	**4.615,47**	. | . *
14	Austria	4.110,48	3.194,12	916,357	. | * .
15	Bulgaria	1.424,83	1.510,83	-85,9932	. * .
16	Denmark	1.675,68	2.736,45	-1 060,77	. * | .
17	Finland	6.763,27	2.412,58	**4.350,69**	. | . *
18	Slovakia	1.236,64	1.689,25	-452,605	. * | .
19	Ireland	487,697	3.052,65	-2.564,95	* . | .
20	Croatia	1.151,90	1.492,50	-340,596	. * | .
21	Lithuania	723,309	1.460,93	-737,620	. * | .
22	Latvia	1.037,81	1.381,87	-344,060	. * | .
23	Slovenia	641,394	1.464,87	-823,474	. * | .
24	Estonia	442,860	1.362,29	-919,426	. * | .
25	Cyprus	197,276	1.337,34	-1.140,06	. * | .
26	Luxembourg	161,030	1.527,15	-1.366,12	. * | .
27	Malta	31,8750	1.294,53	-1.262,65	. * | .
	Sum	101.747,0	101.747,0	0,000	

Conclusions on the significance of the indicators of econometric representation and the assessment of the viability of the econometric model of the final consumption of energy used, renewable and biofuels according to the gross domestic product

The linear one-factor econometric model of final consumption of energy used, renewable and biofuels according to the gross domestic product of the 27 states that make up the European Union, $y = 1{,}232.346+0.004287 \cdot x+u$, is appreciated as a model with reserved viability because not all statistical testing conditions are met to justify a full viability conclusion:

- the correlation ratio (R = 0.911935) has a size very close to the unit, which confirms the existence of a very strong correlation of the final consumption of energy used, renewable and biofuels depending on the gross domestic product. Also, the size of the coefficient of determination (R^2 = 0.831625) offers the possibility to specify that 83.1625% of the change in final consumption of energy used, renewable and biofuels is explained by the change in gross domestic product, the difference up to 100% is the influence residual component or the influence of other factors, not included in the model;

- in the case of the studied correlation, $F_{-statistic}$ = 123.4782 and it is found that this quantity exceeds to a significant extent the tabular value which is 4.24 ($F_{-tabular}$ = 4.24). From the table with the values of the Fisher distribution function is extracted $F_{-tabular}$, which corresponds to a probability of 95% and respectively the number of degrees of freedom, $f_1 = k - 1 = 2 - 1 = 1$ and $f_2 = n - k = 27 - 2 = 25$.

It is thus attested, with sufficient confidence, that the correlation ratio and the coefficient of determination are significantly different from zero or, in other words, the existence of a real correlation between the variables of the studied system is validated.

- the parameters of the regression equation „*a*" and „*b*" are significantly different from zero (for these parameters the null hypothesis is rejected), based on the „*Criterion t*" with significance thresholds of 0.76% and 0.00% respectively, below the maximum limit of 5% (Table 4). Under these conditions, the independent (exogenous) variable, the gross domestic product ensures a significant influence on the final consumption of energy used, renewable and biofuels;

- the econometric model shows, by the size of the regression coefficient „*b*", that an increase by one unit (one million US dollars) of the value of the gross domestic product increases the final consumption of energy used, renewable and biofuels by 0.004287 units (thousand tons of oil equivalent);

- the „*Durbin-Watson statistical coefficient*" (DW = 1.848242) has a size that is in the range 1.469 <1.848242 <4 - 1.469 = 2.531 and we can appreciate with sufficient confidence that the variants of the error term do not autocorrelate, this being a complementary condition to confirm the viability of the regression equation, the model has a correct construction. The conclusion is attested based on the Durbin-Watson distribution for a significance threshold, q = 5%, the number of exogenous variables, k ,= 1 and the number of observations, n = 27 by taking into account the inequality $d_2 < DW < 4 - d_2$;

- the relative expression of the standard error estimate of the regression equation,

$$\hat{V}_{y,\hat{y}} = (1{,}858.682 / 3{,}768.408) \cdot 100 = 49.3227\%$$

provides information that does not support the viability of the model (regression equation) for an extrapolation calculation because it is larger than the acceptable limit of 10%;

- a statistical significance similar to that presented by the estimation of the relative standard error of the regression equation is obtained by calculating and interpreting „*Theil Inequality Coefficient*" (*Th* = 15.9134%). This coefficient can be between zero and one (100) and is considered to be a very good size to assess the viability of the model when it does not exceed the 5% limit;

- the statistical description of the error term series is made by the asymmetry coefficient (Skewness = 0.933389), the vault-flattening coefficient (Kurtosis = 3.796657), the Jarque-Bera statistical coefficient (*J-B* = 4.634458), which follows the distribution law χ^2 with 2 degrees of freedom, and the probability related to the *J-B* coefficient (9.8546%). This information underlies the rejection of the assumption of the error term values according to the law of normal distribution (*conclusion provided by the normality test of the residual variable distribution*), because the probability associated with the *J-B* coefficient is less than the critical limit of 60%. In this situation, the non-confirmation of the hypothesis of normality of the error term distribution, the quality of the parameters of the regression equation to be of maximum probability as well as the calculation of the confidence intervals is obviously affected;

- the heteroscedasticity error test confirms the homoscedasticity property of the final consumption model of energy used, renewable and biofuels, according to the gross domestic product, based on the two statistical criteria applied, „*Criterion F"* and „*Criterion* χ^2 " respectively on the equation of auxiliary regression of the square of the residual levels according to the gross domestic product. Under these conditions, the following assessments can be made:

- the dispersion of errors is constant;

- the application of "*Criterion t*" to verify the significance of the parameters of the regression equation provides fully conclusive statistical information

The size of the final consumption of energy used, renewable and biofuels depending on the size of the gross domestic product is a legitimacy that has been confirmed econometrically as a linear unifactorial model but must be supported both by policy decisions (Mina-Raiu et al., 2021) and by implementing sustainable economic measures to achieve outstanding results, quantified by high efficiency in order to reach the standards of good governance (Raiu, 2015). Achieving this goal is possible only through the simultaneous action of all EU Member States - 27, individually but also through the implementation of coordinated mutual support programs (Mina-

Raiu, 2014), as a necessary and rapid response to environmental degradation, the phenomenon of global warming through pollution.

The results obtained in this field have a transitory value and the support of the final consumption of energy used, renewable and biofuels through the economic development of all the states must be continued with an intensity inscribed on an ascending trend. A decisive role is played by scientific research with an indisputable contribution to the development and improvement of technological solutions with a support inscribed both in the national budgets and in the Union budget.

Gradually reducing the level of pollution and ensuring the health of the living environment is the beneficial solution to the prosperity of all nations.

Selective Bibliography

1. **Andrei, T., Bourbonais, R.** (2008) - „*Econometrics*", Economic Publishing House, Bucharest.
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, Bucharest.
3. **Anghelache, C., Anghel, M.G., Manole, A**. (2015) - "*Economic, financial-banking and informatics modeling"*, Artifex Publishing House, Bucharest.
4. **Mihăilescu, N.** (2021) - „*Statistics and Statistical Basics of Econometrics*", Transversal Publishing House, Bucharest.
5. **Mihăilescu, N.** (2019) - „*Analysis of economic-financial activity - Research methodologies, case studies solved to substantiate economic-financial decisions and knowledge tests*", Transversal Publishing House, Bucharest.
6. **Mihăilescu, N., Căpăţînă, C.** (2018) - „*The reversible impact of the dynamics of the gross domestic product with the imports and exports of goods and services of Romania*", Romanian Statistical Review - Supplement / Nr. 11 and No. 12/2018, Bucharest.
7. **Mina-Raiu, L.** (2014) – "*Public-private partnership impact on local economic development. evidence from Romanian counties*", Economic Sociology, Human Resource Management and Organizational Dynamics, pp. 211-229, Editura ASE, București.
8. **Mina-Raiu, Laura, Bucura, Isabelle-Alexandra, Raiu, Cătălin Valentin** (2021), "*Transposing good practices in the field of Quality Management in Japan, within Romanian Public Administration*", Romanian Statistical Review, nr. 2/2021, ISSN/ eISSN: 1018-046X / 1844-7694, pp. 62-74.
9. **Raiu, Cătălin-Valentin** (2015), „*An Ontology of Good Governance. A Political Theory Approach*." Romanian Journal of Economics 40, no. 1 (49) (2015): 128-143.
10. **Pagliacci, M., Anghelache, G.V., Poca**n, **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, Bucharest.
11. **Stancu, S., Andrei, T., Iacob, A.I., Tușa, E.** (2008) - „*Introduction to econometrics using EViews*", Economic Publishing House, Bucharest.