DOI: 10.53469/jgebf.2024.06(08).06

Application of Regression Analyses and Marginal Analyses for Assessing the Effectiveness of COVID-19 Pandemic Mitigation Measures in Bulgaria

Kuyu John Alfred

University of National and World Economy, Department of National and Regional security, Sofia, "8-mi dekemvri" str. 19
kuyu.john@gmail.com

Abstract: ***Regression analysis is a popular method for determining multiple variables' relationships. We can use it to quantify this dependence between economic measures and the harmful effects of the COVID-19 pandemic. In addition, we can predict change in future periods. The problem with applying the method is the development of hypotheses about the dependence and accuracy of the data. This problem is discussed in this study, as one of the examples we have given is the "real" level of unemployment and the fastened digitalisation of society. On the other hand, we have the definition of the positive and negative effects of dependent variables on measures that are independent variables. We apply marginal analysis to measure the positive effect and determine when a measure is ineffective. The aim and logic of this study are to reduce the "falsification" of the scientific product so that it is based on objective facts and events and not on the feelings and interpretations of the researcher.***

Keywords: Regression analyses, Marginal analyses, Effect, Economic measures, COVID-19

JEL: B40, C12, D60

1. Introduction

This publication was made under the project "Modelling, analysis and optimisation of socio-economic measures to reduce the negative consequences of the COVID-19 pandemic", contract No. KP-06-DK2/2 of 30.03.2021, financed by the Scientific Research Fund" to the Ministry of Education and Science. The content of the publication is the sole responsibility of the author and should in no way be perceived as an expression of the opinion of the "Scientific Research" Fund.

The COVID-19 outbreak is a random event that the standard bell curve does not capture. A pandemic is an event that severely affects the economy of small and medium-sized businesses if we exclude its effects on the health system. The government imposed certain restrictions on the activity of this type of business, which employs the majority of the working population. Accordingly, the government had to make up for its actions by curtailing and sometimes closing businesses. The fact that the study did not find links between economic measures and effects on business does not mean that there are no similar social, political or other ones outside this study's scope.

On the other hand, shortly after the studied period, other unforeseen events took place, such as the armed conflict in Ukraine, economic sanctions and restrictions at the national level, the more active transition of the European Union to the "Green Economy", bankruptcies of speculative organisations in the field of FinTech, Etc. As a result of these events, any subsequent research would be biased. Nevertheless, we are publishing this research with the hope of avoiding any errors.

The main problem in this descent is whether we can prove a link between the economic measures taken concerning the crisis and whether they are related to environmental change. On the other hand, we need to determine if the problems are manageable or systemic and if the changes are due to other factors that we have yet to investigate.

We need to address the context of the problem systematically, as the COVID-19 crisis has effectively changed economic systems at the macro and micro levels. According to systems theory, economics can be defined as a complex, unmanageable system.

Systems theory is a cross-disciplinary framework for comprehending intricate systems in nature, society, and technology. It emphasises the interconnections and interdependence between the different components of a system rather than focusing on individual parts in isolation. The theory is based on concepts from various fields, including biology, engineering, physics, and sociology, to analyse and explain the behaviour of systems.

One key concept of systems theory is that systems are composed of interrelated elements that work together to achieve a common goal. These elements can be physical components, such as parts of a machine, or abstract objects, such as ideas or concepts. Systems theory also recognises that systems are dynamic and constantly changing, with feedback loops that allow self-correction and adaptation.

Another critical aspect of systems theory is the idea of system boundaries. We define systems by their boundaries, which mark the system's boundaries and distinguish it from its environment. Understanding the boundaries of a system

is crucial to understanding how it interacts with other systems and the environment.

Systems theory is an essential concept with far-reaching applications in various fields, such as ecology, business management, and social psychology. It is the key to understanding complex phenomena such as the behaviour of ecosystems, the functioning of organisations, and the dynamics of social groups. In practice, systems theory is an indispensable tool for identifying design problems and solutions that consider the intricate interdependencies between various system elements. The tasks we have set ourselves are related to establishing methodology, model, and analysis.

2. Regression analyses, Dependent variables / Indicators

Why Regression analyses?

- It can be an objective analysis.
- It is a grey approach if appropriately used.
- By regression analyses, we can prove some form of correlation between the variables.
- We can measure the quantity of measures, and we can measure time.
- If we do not have several measures, for example (How much money is given for digitalisation?), we can still measure the dependence by the time of application.

Regression analysis is a set of statistical techniques used to evaluate the relationship among variables. A common misconception is that regression analysis is the same as correlation analysis. Correlation analysis measures the association between two variables, whereas regression analysis focuses on understanding if one or more variables can explain some process. More specifically, regression analysis determines if one or more independent variables can explain the variability in a dependent variable. We need to consider which is our dependent variable and which are our independent variables.[1]

There are four broad classes of applications of regression analysis.

Descriptive or explanatory. Descriptive analyses involve the analysis of single-variable statistics and univariable analyses. These analyses are helpful for quickly identifying notably extreme data or outliers and obtaining p-values[2]. Predictive - As predictive analytics is a tool for machine learning and big data, regression modelling is a tool for predictive analytics—one of the primary tools. Regression analysis entails looking at dependent variables (outcomes) and an independent variable (the action) while assessing the association's strength. In other words, it looks to understand if there is a relationship between variables and how strong that relationship is.[3]

We can use various criteria when comparing different regression or time series forecasting models for a given dataset. These criteria include various error measures in the estimation period, such as root mean squared error, mean absolute error, mean absolute percentage error, mean absolute scaled error, mean error, and mean percentage error. If we have done out-of-sample testing, we can also consider error measures in the validation period.

Residual diagnostics and goodness-of-fit tests: plots of actual and predicted values; plots of residuals versus time, versus predicted values, and other variables; residual autocorrelation plots, cross-correlation plots, and tests for normally distributed errors; measures of extreme or influential observations; tests for excessive runs, changes in mean, or changes in variance (lots of things that can be "OK" or "not OK")

Qualitative considerations: intuitive reasonableness of the model, simplicity of the model, and usefulness for decision making![4]

Regression analysis is a statistical method used to determine the strength and direction of the relationship between variables. It answers the question: Which factors matter most? Which can we ignore? How do those factors interact with one another? Moreover, most importantly, how certain are we about all these factors?

In regression analysis, we define variables as factors. In Redman's example, we try to understand or predict the dependent variable, such as monthly sales. Independent variables are those that affect the dependent variable.[5]

The main variables that we use to prove an economic effect are:

- GDP
- Employment/Unemployment
- PPP (Purchasing power parity)
- Inflation/Deflation
- Digitalisation
- Supply Chain
- Demographics (including migration)
- Exchange rate
- Economic growth
- International investments
- Others positive or negative

Accordingly, we must determine whether the variables are dependent or independent and compare them with the value of the measures and the change in price values.

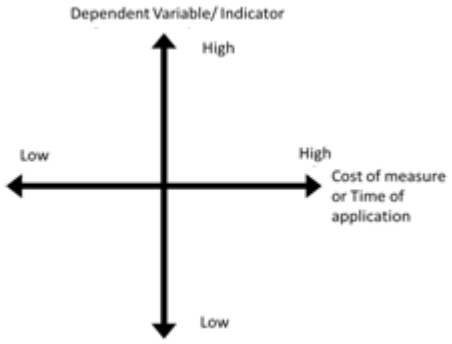

Figure 1: Baseline model for variable distribution

Hypothesis Conditions

Accordingly, we need to determine whether the indicator has a positive or negative value and importance for the economy.

- To perform regression analyses, we need data, models, and time. Instruments are optional and valuable. (I prefer MS Excel Pro.)
- Some indicators have a positive and negative meaning. Others have neutral. Some have meaning depending on the context.
- The positive and negative are also discussible.

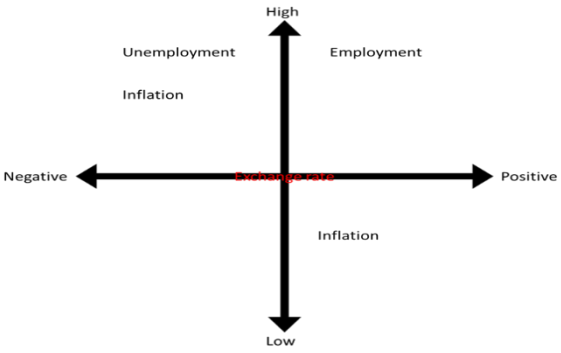

Figure 2: Example of distribution of indicators

First step Hypothesis

A high unemployment rate is strongly negative, while a low one is strongly positive.

H0 – There is a connection between the rate of unemployment and measures

H1 – There is no connection between the rate of unemployment and measures

Second step forecasts

Condition for data: We can Forecast only up to 1/3 of the time for which we have data

We can measure the margin between the reported and most expected value by the forecast.

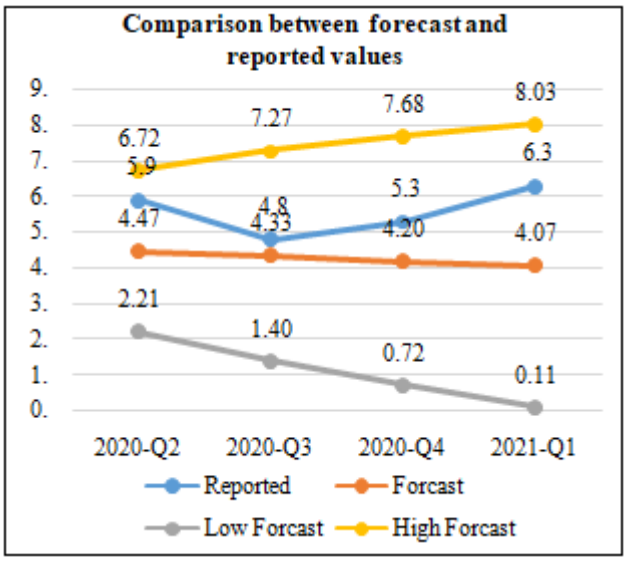

Figure 3: Comparison between unemployment data and projected unemployment data during the COVID-19 crisis

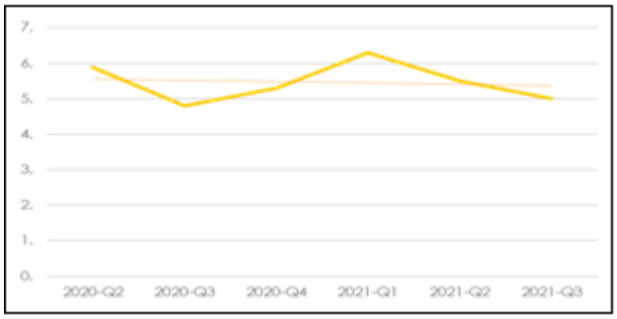

Figure 4: Regression Data/ Measures/Unemployment

Third step: Regression Analyses

- Y – unemployment
- X1 – total average measures (I have taken the total costs of the measures and distributed them average by the YQ (486,55 million per Q))
- Notes: Not very correct; Need more data on the exact Expenses.

Coefficients	
5,466667	
0	
F	Significance F
2,655306834	0,178538
R Square	0,398976
Adjusted R Square	0,24872

Third step: Regression Analyses

Y – unemployment

X1 – total average 60/40(166,6667 m)

X2 – Total average OPIC grant scheme for SMEs (45,55 m)

Notes: Not very correct; Need more data on the exact Expenses.

R Square	0
Adjusted R Square	-0,4

F	Significance F
0	1

	Coefficients
Intercept	5,466667
X Variable 1	0
X Variable 2	0

Third step: Regression Analyses

Y – Employment

X1 – total average measures (I have taken the total costs of the measures and distributed them average by the YQ (486,55 million per Q))

X2 – total average 60/40(166,6667 m)

X3 – Total average OPIC grant scheme for SMEs (45,55 m)

Notes: Not very correct; Need more data on the exact Expenses.

Regression Statistics	
Multiple R	0,75912
R Square	0,576263
Adjusted R Square	-0,02967
Standard Error	5,050993
Observations	6

F	Significance F
5,439807	0,159206

	Coefficients
	2976,167

Example Data

Linear regression is an essential statistical tool researchers use to develop and evaluate predictive models. While it can accurately determine the strength of a group of potential

predictors in forecasting a relevant criterion, it needs to identify the most crucial predictor variables. Various established techniques exist to establish variable importance, but none are considered definitive or indisputable.[6]

We calculate the unemployment forecast with the forecast function in MS EXCEL.

Table 1: Unemployment data and forecast hypothesis[7]

Data Eurostat		Forecast 2020-2021 without covid		
Weather	Unemployment%	Weather	Period	Unemployment%
2000-Q1	18,7	**2000-Q1**	1	18,7
2000-Q2	16,2	**2000-Q2**	2	16,2
2000-Q3	16,3	**2000-Q3**	3	16,3
2000-Q4	16,6	**2000-Q4**	4	16,6
2001-Q1	22,3	**2001-Q1**	5	22,3
2001-Q2	19,9	**2001-Q2**	6	19,9
2001-Q3	19,1	**2001-Q3**	7	19,1
2001-Q4	20,0	**2001-Q4**	8	20,0
2002-Q1	20,0	**2002-Q1**	9	20,0
2002-Q2	18,1	**2002-Q2**	10	18,1
2002-Q3	17,6	**2002-Q3**	11	17,6
2002-Q4	17,1	**2002-Q4**	12	17,1
2003-Q1	15,6	**2003-Q1**	13	15,6
2003-Q2	13,8	**2003-Q2**	14	13,8
2003-Q3	12,7	**2003-Q3**	15	12,7
2003-Q4	12,8	**2003-Q4**	16	12,8
2004-Q1	13,4	**2004-Q1**	17	13,4
2004-Q2	12,1	**2004-Q2**	18	12,1
2004-Q3	11,0	**2004-Q3**	19	11,0
2004-Q4	11,9	**2004-Q4**	20	11,9
2005-Q1	11,4	**2005-Q1**	21	11,4
2005-Q2	10,0	**2005-Q2**	22	10,0
2005-Q3	9,2	**2005-Q3**	23	9,2
2005-Q4	9,9	**2005-Q4**	24	9,9
2006-Q1	9,7	**2006-Q1**	25	9,7
2006-Q2	9,0	**2006-Q2**	26	9,0
2006-Q3	8,9	**2006-Q3**	27	8,9
2006-Q4	8,4	**2006-Q4**	28	8,4
2007-Q1	8,0	**2007-Q1**	29	8,0
2007-Q2	6,8	**2007-Q2**	30	6,8
2007-Q3	6,6	**2007-Q3**	31	6,6
2007-Q4	6,1	**2007-Q4**	32	6,1
2008-Q1	6,5	**2008-Q1**	33	6,5
2008-Q2	5,8	**2008-Q2**	34	5,8
2008-Q3	5,1	**2008-Q3**	35	5,1
2008-Q4	5,0	**2008-Q4**	36	5,0
2009-Q1	6,4	**2009-Q1**	37	6,4
2009-Q2	6,3	**2009-Q2**	38	6,3
2009-Q3	6,7	**2009-Q3**	39	6,7
2009-Q4	7,9	**2009-Q4**	40	7,9
2010-Q1	10,2	**2010-Q1**	41	10,2
2010-Q2	10,0	**2010-Q2**	42	10,0
2010-Q3	9,6	**2010-Q3**	43	9,6
2010-Q4	11,4	**2010-Q4**	44	11,4
2011-Q1	12,2	**2011-Q1**	45	12,2
2011-Q2	11,2	**2011-Q2**	46	11,2
2011-Q3	10,3	**2011-Q3**	47	10,3
2011-Q4	11,4	**2011-Q4**	48	11,4
2012-Q1	12,9	**2012-Q1**	49	12,9
2012-Q2	12,3	**2012-Q2**	50	12,3
2012-Q3	11,5	**2012-Q3**	51	11,5
2012-Q4	12,4	**2012-Q4**	52	12,4
2013-Q1	13,8	**2013-Q1**	53	13,8
2013-Q2	13,0	**2013-Q2**	54	13,0
2013-Q3	12,0	**2013-Q3**	55	12,0
2013-Q4	13,1	**2013-Q4**	56	13,1
2014-Q1	13,0	**2014-Q1**	57	13,0
2014-Q2	11,4	**2014-Q2**	58	11,4
2014-Q3	10,8	**2014-Q3**	59	10,8
2014-Q4	10,6	**2014-Q4**	60	10,6
2015-Q1	10,6	**2015-Q1**	61	10,6
2015-Q2	9,9	**2015-Q2**	62	9,9
2015-Q3	8,3	**2015-Q3**	63	8,3
2015-Q4	7,9	**2015-Q4**	64	7,9
2016-Q1	8,6	**2016-Q1**	65	8,6
2016-Q2	8,1	**2016-Q2**	66	8,1
2016-Q3	7,0	**2016-Q3**	67	7,0
2016-Q4	6,7	**2016-Q4**	68	6,7
2017-Q1	6,9	**2017-Q1**	69	6,9
2017-Q2	6,3	**2017-Q2**	70	6,3
2017-Q3	5,9	**2017-Q3**	71	5,9
2017-Q4	5,6	**2017-Q4**	72	5,6
2018-Q1	5,7	**2018-Q1**	73	5,7
2018-Q2	5,5	**2018-Q2**	74	5,5
2018-Q3	5,0	**2018-Q3**	75	5,0
2018-Q4	4,7	**2018-Q4**	76	4,7
2019-Q1	5,0	**2019-Q1**	77	5,0
2019-Q2	4,2	**2019-Q2**	78	4,2
2019-Q3	3,7	**2019-Q3**	79	3,7
2019-Q4	4,1	**2019-Q4**	80	4,1
2020-Q1	4,6	**2020-Q1**	81	4,6
2020-Q2	5,9	**2020-Q2**	82	
2020-Q3	4,8	**2020-Q3**	83	
2020-Q4	5,3	**2020-Q4**	84	
2021-Q1	6,3	**2021-Q1**	85	
2021-Q2	5,5			
2021-Q3	5,			

Table 2: Mean for the regression

2012-Q1	12,9	2013-Q1	13,8	2014-Q1	13,0	2015-Q1	10,6
2012-Q2	12,3	2013-Q2	13,0	2014-Q2	11,4	2015-Q2	9,9
2012-Q3	11,5	2013-Q3	12,0	2014-Q3	10,8	2015-Q3	8,3
2012-Q4	12,4	2013-Q4	13,1	2014-Q4	10,6	2015-Q4	7,9
	12,275		12,975		11,45		9,175
2016-Q1	8,6	2017-Q1	6,9	2018-Q1	5,7	2019-Q1	5,0
2016-Q2	8,1	2017-Q2	6,3	2018-Q2	5,5	2019-Q2	4,2
2016-Q3	7,0	2017-Q3	5,9	2018-Q3	5,0	2019-Q3	3,7
2016-Q4	6,7	2017-Q4	5,6	2018-Q4	4,7	2019-Q4	4,1
	7,6		6,175		5,225		4,25
2020-Q1	4,6						
2020-Q2	5,9						
2020-Q3	4,8						

2020-Q4	5,3						
	5,15						
2021-Q1	6,3						

Table 3: Average for unemployment and migration

	Average unemployment	Migration
	And	X
2012	12,275	-8676
2013	12,975	-11354
2014	11,45	-14347
2015	9,175	-13765
2016	7,6	-16541
2017	6,175	-13932
2018	5,225	-15094
2019	4,25	-14376
2020	5,15	20394

Table 4: Unemployment measures

Measure 60/40	OPIC grant scheme for SMEs
1000	273,3
166,6666667	45,55

3. Problems with data and example

We tested over 70 regression-based models during the project's first six months. We base the example on unemployment and measures due to the availability of the most data. This model comes closest to some form of positive outcome.

Table 5: List of measures

At the expense of the budget:	MILLION. LV
Prophylactic and anti-epidemic measures (disinfectants, barriers, masks) under PMS 40, thermal cameras	14,6
Hot lunch	5,2
Insulation control measures	9,6
For overtime work of police officers	161,4
For the participation of the service members (not applied)	168,1
Posting serves the RHC and RHI	0,5
Replenishment of the state reserve	4
Distance learning of students	2,2
Increase in the capital of BDB	700
Measure 60/40	1000
At the expense of operational programs:	0
Protection of the population from threats to public health (remuneration, business trips, and materials) 1,000 for 14,000 doctors, nurses, nurses, and laboratory assistants, HRD OP	60
Patronage care for adults and people with disabilities OPCHR	45
term employment schemes	254,8
OPIC grant scheme for SMEs	273,3
OPIC financial instrument	200
Medical and diagnostic equipment (ventilation devices, tests, Etc.) OPRD	40,4
Total	2919,3

Table 6: Unemployment – measures hypothesis

	And	X1	X2	X3
2020-Q2	5,9	486,55	166,6667	45,55
2020-Q3	4,8	486,55	166,6667	45,55
2020-Q4	5,3	486,55	166,6667	45,55
2021-Q1	6,3	486,55	166,6667	45,55
2021-Q2	5,5	486,55	166,6667	45,55
2021-Q3	5,	486,55	166,6667	45,55

Summary Output

Regression Statistics	
Multiple R	0,759119567
R Square	0,576262518
Adjusted R Square	-0,029671853
Standard Error	5,050992732
Observations	6

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	138,783223	46,26107	5,439806879	0,159206196
Residual	4	102,0501103	25,51253		
Total	7	240,8333333			

	Coefficients	*Standard Error*	*t Stat*
Intercept	2976,166667	2,062059148	1443,298
X1	0	0	65535
X2	0	0	65535
X3	0	0	65535
P-value	*Lower 95%*	*Upper 95%*	*Lower 95,0%*
1.38269E-12	2970,441473	2981,891861	2970,441473
#NUM!	0	0	0
#NUM!	0	0	0
#NUM!	0	0	0

Upper 95,0%
2981,891861
0
0
0

Residual Output

Observation	*Predicted Y*
1	2976,166667
2	2976,166667
3	2976,166667
4	2976,166667
5	2976,166667
6	2976,166667
Residuals	*Standard Residuals*
-10,16666667	-1,73328
8,833333333	1,505967
1,833333333	0,312559
-4,166666667	-0,71036
-2,166666667	-0,36939
5,833333333	0,994506

Probability Output

Percentile	*And*
8,333333333	2966
25	2972
41,66666667	2974
58,33333333	2978
75	2982
91,66666667	2985

Regression Unemployment — Migration

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0,21885906
R Square	0,047899288
Adjusted R Square	-0,088115099
Standard Error	3,481424655
Observations	9

ANOVA		
	df	*SS*
Regression	1	4,268332179
Residual	7	84,84222338
Total	8	89,11055556
MS	*F*	*Significance F*
4,268332179	0,352163	0,571559398
12,12031763		

	Coefficients	*Standard Error*
Intercept	7,635425211	1,558505303
X Variable 1	-6.33608E-05	0,00010677
t Stat	*P-value*	*Lower 95%*
4,899197452	0,001755	3,950145775
-0,593433562	0,571559	-0,000315831
Upper 95%	*Lower 95,0%*	*Upper 95,0%*
11,32070465	3,950145775	11,32070465
0,00018911	-0,000315831	0,00018911

RESIDUAL OUTPUT	
Observation	*Predicted Y*
1	8,185143629
2	8,354823888
3	8,544462803
4	8,507586809
5	8,683476428
6	8,518168065
7	8,591793331
8	8,546300266
9	6,34324478
Residuals	Standard Residuals
4,089856371	1,255876776
4,620176112	1,418722653
2,905537197	0,892206561
0,667413191	0,204943316
-1,083476428	-0,33270432
-2,343168065	-0,719519242
-3,366793331	-1,033844999
-4,296300266	-1,319269734
-1,19324478	-0,366411011

4. Conclusion and Future Work - Marginal Analyses

Data problems

- Different data from various sources.
- Absolute versus relative values.
- For example, the data for unemployment needs to be corrected. It measures only registered the registered unemployed in the labour offices.
- All cases have a correlation cases correlate, but the p-value does not fit.
- There are background factors that require more study.
- Which costs are connected to the pandemic?
- What can the aid be spent?
- Time of application of the measure and time of receiving the aid.

Marginal analyses example

Table 7: MA example

TIME		Marginal	AVG Total Expenses	Average total 60/40	AVG tot. Costs.	AVG 60/40 PP
2019-Q2	3 115					
2019-Q3	3 082	-33				
2019-Q4	3 065	-17				
2020-Q1	3 029	-36				
2020-Q2	2 966	-63	486,55	166,6667	0,164042481	0,056192414
2020-Q3	2 985	19	486,55	166,6667	0,162998325	0,05583474
2020-Q4	2 978	-7	486,55	166,6667	0,163381464	0,055965984
2021-Q1	2 972	-6	486,55	166,6667	0,163711306	0,05607897
2021-Q2	2 974	2	486,55	166,6667	0,16360121	0,056041258
2021-Q3	2 982	8	486,55	166,6667	0,163162307	0,055890912

5. Conclusion and Future Studies

- If we are capable of building the right hypothesis, we can prove or reject the correlation between the variables;
- By regression, we can show the objective importance of a measure;
- By marginal analyses, we can measure the marginal cost and marginal effect;
- For the future study, we will implement the cognitive fuzzy models. Fuzzy models will allow the definition of structural problems in economics.

References

[1] Geog 418/518, 'REGRESSION ANALYSIS', ArcGIS StoryMaps. Accessed: Apr. 26, 2023. [Online]. Available: https://storymaps.arcgis.com/stories/6b694548e67d47ab99306d8fe68910f8

[2] Medistica, 'pvalue.io, a GUI of R statistical software for scientific medical publications.', pvalue.io. Accessed: Apr. 26, 2023. [Online]. Available: https://www.pvalue.io

[3] M. Rudat, 'Predictive Analytics and Regression Models Explained', GutCheck. Accessed: Apr. 26, 2023. [Online]. Available: https://gutcheckit.com/blog/predictive-analytics-regression-models-explained/

[4] 'How to compare regression models'. Accessed: Apr. 26, 2023. [Online]. Available: https://people.duke.edu/~rnau/compare.htm

[5] A. Gallo, 'A Refresher on Regression Analysis', *Harvard Business Review*, Nov. 04, 2015. Accessed: Apr. 26, 2023. [Online]. Available: https://hbr.org/2015/11/a-refresher-on-regression-analysis

[6] M. T. Braun and F. L. Oswald, 'Exploratory regression analysis: A tool for selecting models and determining predictor importance', *Behav Res*, vol. 43, no. 2, pp. 331–339, Jun. 2011, doi: 10.3758/s13428-010-0046-8.

[7] 'Statistics | Eurostat'. Accessed: Apr. 26, 2023. [Online]. Available: https://ec.europa.eu/eurostat/databrowser/view/UNE_RT_Q__custom_2395064/default/table?lang=en