



Correlational-regression Analysis Application for the Forecast of the Specialists with Higher Education Requirement in Russian Economy

Nina Potekhina^{1*}, Yulia Shulinina², Natalia Kuzmina³, Ludmila Potalisina⁴, Inessa Sannikova⁵

¹Department of Management, National Research Tomsk Polytechnic University, Tomsk, Russia, ²Department of Management, National Research Tomsk Polytechnic University, Tomsk, Russia, ³Department of Management, National Research Tomsk Polytechnic University, Tomsk, Russia, ⁴Department of Economics, Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russia, ⁵Department of Management, National Research Tomsk Polytechnic University, Tomsk, Russia.
*Email: potechinanina@gmail.com

ABSTRACT

The present study was intended to investigate a hypothesis about the impact of the following parameters: Dynamics of gross domestic product, dynamics of fixed assets, dynamics of labor productivity, dynamics of the level of remuneration and dynamics of fixed asset investments on the number of employees with higher education in Russia. The correlational-regression analysis confirmed the influence of dynamics of fixed assets on the number of employees with higher education. Also authors have generated the forecast about future demand for specialists with higher education and dynamics of fixed assets for the period from 2015 to 2025 years.

Keywords: Correlation Analysis, Regression Analysis, Forecast of the Specialists with Higher Education

JEL Classifications: C100, J240

1. INTRODUCTION

Workforce is the primary factor determining successful economic development in the modern world. Labor of employees with higher educational level is characterized by higher capacity. From Adam Smith to Becker, economic thinkers have highlighted the importance of labor force. Becker showed the importance of higher education as part of human capital to increase personal income, labor productivity and gross domestic product (GDP) (Becker, 1993). Moreover, many studies report that, there is a link between human capital and economic development (Schultz, 1993; Hanushek, 2013; Pelinescu, 2015). Therefore, the question about the demand in workers with high qualification stands important in economy. We have used the method of correlational-regression analysis to forecast the demand in specialists with higher education to economy of Russia. This method considers accentuation of the two key components: Correlational analysis and regression analysis. When correlational analysis gives determination

of correlation colligation and direction between researched parameters, regression analysis helps to determine a type of functional dependence between them. Joint application of these methods of analysis allows us to:

- Create the correlational-regression model, which describes the connection between researched parameters.
- Make a conclusion about the impact of a parameter on a resulting sign.
- Forecast a change of a resulting sign in time.

It is evident, that this type of analysis is very promising for the economic theory (Popescu, 2015; Maximova, 2011).

Authors performed the correlational-regression analysis to prove or disprove the hypothesis about the impact on number of employed people of the following parameters: Dynamics of GDP, dynamics of fixed assets, dynamics of labor productivity, dynamics of the level of remuneration, and dynamics of fixed asset investments in Russia.

2. THE CORRELATIONAL-REGRESSION ANALYSIS IMPLEMENTATION METHODOLOGY

The correlational-regression analysis implementation methodology considers accentuation of several stages. There are several points of view regarding the number of stages. The methodology of the correlational-regression analysis is described in different studies (Gujarati, 2003; Eliseeva et al., 2007; Weisberg, 2005).

Firstly, we need to choose the factors, which could affect a resulting sign. On the second stage, we perform the correlational analysis. As a result, we have a matrix of paired correlation coefficients, which is used as the basis for analysis of parameters interdependence. Then we select parameters, mostly affecting

Table 1: Basic data to perform the correlational-regression analysis

Years	Y	X ₁	X ₂	X ₃	X ₄	X ₅
2004	16762.33	107.2	106.5	110.6	101.6	113.7
2005	17904.82	106.4	105.5	112.6	101.9	110.2
2006	18952.23	108.2	107.5	113.3	102.4	116.7
2007	20735.7	108.5	107.5	117.2	103.1	122.7
2008	20945.9	105.2	104.8	111.5	103.6	109.9
2009	19573.75	92.2	95.9	96.5	103.2	86.5
2010	20350.71	104.5	103.2	105.2	103	106.3
2011	21115.27	104.3	103.8	102.8	104	110.8
2012	21749.81	103.4	103.0	108.4	104.3	106.8
2013	22631.09	101.3	101.9	104.8	104.1	100.8
2014	23035.56	100.6	100.8	101.2	103.8	97.3

Table 2: Correlational analysis results

	Y	X ₁	X ₂	X ₃	X ₄	X ₅
Y	1					
X ₁	-0.305	1				
X ₂	-0.375	0.988	1			
X ₃	-0.361	0.886	0.918	1		
X ₄	0.928	-0.395	-0.451	-0.437	1	
X ₅	-0.313	0.961	0.974	0.901	-0.354	1

Table 3: Results of regression analysis

Regression statistics						
Multiple R	0.946					
R ²	0.895					
Standardized R ²	0.790					
Standard error	880.092					
Observations	11.000					
Dispersion analysis						
	df	SS	MS	F	Significance of F	
Regression	5.000	32997554.967	6599510.99	8.520	0.017	
Remainder	5.000	3872813.479	774562.69			
In total	10.000	36870368.446				
	Coefficients	Standard error	t-statistic	P-value	Bottom 95%	Top 95%
Y-crossing	-203145.328	76227.504	-2.665	0.045	-399094.36	-7196.291
X ₁	328.468	464.538	0.707	0.511	-865.665	1522.600
X ₂	-164.130	920.800	-0.178	0.866	-2531.122	2202.863
X ₃	50.844	122.258	0.416	0.695	-263.431	365.119
X ₄	2067.915	425.622	4.859	0.005	973.818	3162.013
X ₅	-115.773	142.129	-0.815	0.452	-481.127	249.581

the resulting sign, and exclude from the system the parameters with a weak connection with it. Besides, we discard parameters with a high paired correlation coefficient. On the second stage we perform the regression analysis. Based on its results we analyze the significance of the determination coefficient, F Fisher criterion, compare regression coefficients with a standard error to Student's t-test, and create a regression equation.

The valuation method used is the following equation:

$$y = a_0 + a_1 \times X_1 + a_2 \times X_2 + \dots + a_n \times X_n \tag{1}$$

a₀ - Constant term

a₁ - Regression coefficient.

3. IMPLEMENTATION OF CORRELATIONAL-REGRESSION ANALYSIS

We research the number employed people in economy of Russia, having higher education (Y), as the resulting sign, dynamics of GDP (X₁), dynamics of labor productivity (X₂), dynamics of the level of remuneration (X₃), dynamics of fixed assets, and dynamics of fixed asset investments in Russia (X₅) - As factor features. Data for implementation of correlational-regression analysis are taken from materials of the Federal State Statistics Service and presented in the Table 1 (Russia in Figures, 2015; Statistical Yearbook of Russia, 2010, 2014).

For convenience and time saving we use the software product Microsoft Excel, add-in "Analysis ToolPak," functions "Correlation," "Regression." The results are presented in Tables 2 and 3.

The paired correlation coefficients received based on the results of calculations are in the range from -1 to 1. If the value of a coefficient tends to "1" or "-1," we can say, that there is a close connection between the selected parameters. If the value of a

Table 4: Results of repeat regression analysis

Regression statistics						
Multiple R	0.928					
R ²	0.861					
Standardized R ²	0.846					
Standard error	754.672					
Observations	11					
Dispersion analysis						
	df	SS	MS	F	Significance of F	
Regression	1	31744595.03	31744595.03	55.738	3.828	
Remainder	9	5125773.414	569530.379			
In total	10	36870368.45				
	Coefficients	Standard error	t-statistic	P-value	Bottom 95%	Top 95%
Y-crossing	-183718.176	27333.543	-6.721	8.641	-245550.946	-121885.407
Variable X_4	1977.671	264.897	7.466	3.828	1378.432	2576.911

Table 5: Forecast data

Years	The forecast of the number of employed people in economy, having higher education in thousand persons	The forecast of availability of fixed assets by the end of the year
2015	23221.77	104.64
2016	23701.80	104.88
2017	24181.84	105.12
2018	24661.87	105.37
2019	25141.91	105.61
2020	25621.94	105.85
2021	26101.98	106.09
2022	26582.01	106.34
2023	27062.05	106.58
2024	27542.08	106.82
2025	28022.12	107.07

coefficient is laid from “-1” to “0,” the connection between selected parameters is inverse.

The degree of colligation between selected parameters is estimated by the Chaddock scale. The received values of linear paired correlation coefficients indicate a moderate connection of the number of the employed people with the higher education with selected parameters besides a parameter - Index of availability of fixed assets by the end of the year, having the strong connection with the resulting sign.

The received value of the determination coefficient equal to 0.89 shows 89% of community variation of employed people with higher education is explained by variation of factor parameters and 11% - By parameters not researched in this model. The calculated F Fisher criterion equal to 8.52 exceeds table F - 5.12. Therefore, the received dependence is statistically significant. The parameters, absolute regression coefficient value of which is greater than standard error are excluded from analysis. Comparative analysis of received values of regression and standard error coefficients revealed significant only the index of availability of fixed assets by the end of the year. The calculated values of Student's t-test show, that a_0 and a_4 parameters are statistically significant and a_1 , a_2 , a_3 and a_5 parameters have been formed under influence of random causes; therefore, factors can be excluded from an equation as not definitive. This fact is also

proved by P-value: Where the significance level of $P < 0.05$ level we accepted.

After an exclusion of secondary parameters, we have repeated the regression analysis presented in Table 4.

The received value of the coefficient of determination, F Fisher criterion, comparative analysis of values of regression and standard error coefficients, and the Student's t-test value show that the received dependence is statistically significant.

The verification of the model for an adequacy, prognostic suitability, and exceptions of unnecessary parameters gave the following dependence of the number of employed people in economy having higher education in Russia:

$$Y = -183718.17 + 1977.67 \times X_4 \quad (2)$$

Y - Number of employed people in economy having higher education

X_4 - Index of availability of fixed assets by the end of the year.

4. A FORECAST OF A DEMAND OF SPECIALISTS WITH HIGHER EDUCATION

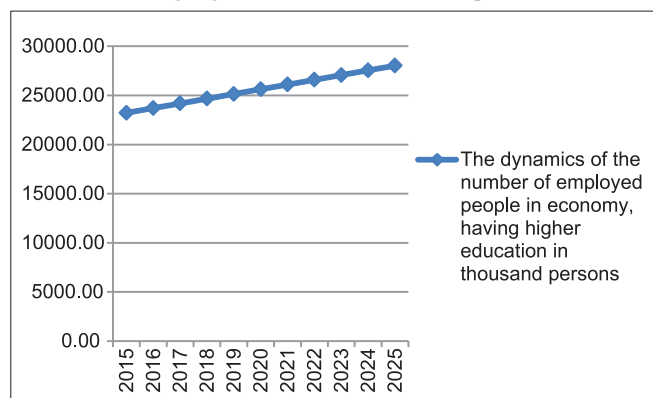
On basis of correlational-regression analysis, the authors have predicted the index of availability of fixed assets by the end of the year, after predicted the number of employed people in economy, having higher education. To this was applied a linear regression equation. The extrapolation of series to dynamics was used for it. The results of forecasting are presented in Table 5 and Figure 1.

The calculations showed that the demand of specialists with higher education for economy of Russia has a stable trend to growth. In general, in accordance with a forecast the number of employed people in economy, having higher education will increase by 2025 by 21.6%.

5. CONCLUSION

The presented model needs to be more developed in future, because during the correlational-regression analysis the

Figure 1: The forecast of the number of employed people in economy, having higher education in thousand persons



following factors were excluded: Dynamics of GDP, dynamics of labour productivity, dynamics of the level of remuneration, and dynamics of fixed asset investments in Russia, that signalize of other factors not taken into consideration by authors, influencing the resulting sign existence. Besides, during the research authors proved, that correlational-regression analysis application for forecasting of the number of employed people with higher education is valid only in the short and medium terms. In the long term, the demand of specialists with higher education could be affected by crisis phenomena in economy, demographic trends, change of scientific-technological systems, and reinforcement of global processes in the world and other. Also, study of this model, from the point of view on various specialties, seems to be interesting, that would provide more exact determination of

the demand of specialists with higher education for particular branches of economy.

REFERENCES

- Becker, G.S. (1993), *Human capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. USA: University of Chicago Press.
- Eliseeva, I., Kuryshcheva, S., Kasteeva, T. (2007), *Econometrics*. Moscow: Finance and Statistics.
- Gujarati, D.N. (2003), *Basic Econometrics*. International Edition. New York: McGraw-Hill.
- Hanushek, E.A. (2013), Economic growth in developing countries: The role of human capital. *Economics of Education Review*, 37, 204-212.
- Maxsimova, A. (2011), Probabilistic model need for specialists with higher professional education for economy of Tomsk region. *Bulletin of Tomsk State Pedagogical University*, 12, 87-92.
- Pelinescu, E. (2015), The impact of human capital on economic growth. *Procedia Economics and Finance*, 22, 184-190.
- Popescu, A. (2015), Analysis of the Dynamics of the Gross Domestic Product and its Factors of Influence in Romania's Agriculture (Conference Paper). In: *Proceedings of the 25th International Business Information Management Association Conference - Innovation Vision 2020: From Regional Development Sustainability to Global Economic Growth*, IBIMA, 1379-1393.
- Russia in Figures. (2015). Moscow: Rosstat
- Schultz, T.W. (1993), The economic importance of human capital in modernization. *Education Economics*, 1, 13-19.
- Statistical Yearbook of Russia. (2010). Moscow: Rosstat
- Statistical Yearbook of Russia. (2014). Moscow: Rosstat
- Weisberg, S. (2005), *Applied Linear Regression*. New Jersey: John Wiley & Sons, Inc.