

Modeling the Relationship of Social Product and Indicators of Demography, Labor, and Employment as a Scientific Basis for the Digital Economy of the Region

Kozlova E.I.¹ Novak M.A.^{1,*} Karlova M.Yu.²

¹*Department of Economics, Lipetsk State Technical University, Lipetsk, Russia*

²*Department of Mathematics and Physics, Lipetsk State Pedagogical University, Lipetsk, Russia*

*Corresponding author. Email: ferz235@mail.ru

ABSTRACT

The main resource of the digital economy is information that, thanks to digital trends, has no limits of use [1]. Economic information, clothed in the form of science-based models of planning and management, in a digital economy makes it possible to increase the level of forecasting by the development of socio-economic systems [2]. The study aims to identify causal relationships in the economy of a particular region between a social product and such socio-economic indicators as demography, labor and employment, based on correlation and regression analysis. The article analyzes the degree of dependence and tightness of the relationship (correlation) between GRP and the main indicators of the development of the Lipetsk region. Trend forecasting models and multiple regression equations are constructed, and the significance of the obtained dependencies is estimated. The results obtained in this study provide an understanding of the role of the number of employed in the dynamics of the gross regional product of the Lipetsk region.

Keywords: *gross regional product, demographic situation, labor resources, average annual number of employees, correlation and regression analysis, trend analysis, multiple regression models, forecast*

1. INTRODUCTION

The slowdown in the growth rate of the Russian economy, which is estimated at the regional level as the growth rate of gross regional product (GRP), entails the need for strategic planning of regional socio-economic systems. Historically significant tool for studying economic mechanisms in micro- and macroeconomics are economic-mathematical models [3]. Economic models (one leveling or multi-parameter) allows you to see forecast errors and improve understanding of the economy [4]. Digitalization of the economy and the introduction of information technology create opportunities for a more complete reflection in economic models of causal relationships, streamlining information flows and improving the quality of management decisions [5].

The absolute value of GRP acts as an objective indicator reflecting the contribution of the region to the country's economy, as well as an important tool for studying the structural and sectoral proportions of the mesoscale of the economy [6]. In the framework of another approach, GRP is the main indicator of the productivity of the socio-economic system of the region [7].

Regional development means a progressive change in the internal socio-economic structure of the region. It is determined by the action of a whole complex of social

factors, including socio-demographic, administrative, territorial-geographical and other factors [8, p. 669].

The demographic prerequisites for the economic development of the region include, first of all, the natural reproduction of the population, population, quality of labor resources. The specific characteristics of these parameters depend on the stage of demographic development inherent in a particular region. The state of demographic processes in a certain region at a certain time is characterized by the term demographic situation [9].

The problem of the influence of demographic conditions on the country's economic indicators has been at the center of discussions of foreign researchers for many decades. For example, David E. Bloom and Richard B. Freeman, exploring the relationship between population growth and economic growth in developing countries from 1965-1985, concluded that “the timing and components of population growth are important elements in the process of economic development” [10]. James A. Brander and Steve Dowrick based on a data set from 107 country groups for 1960-85. found that “the decline in fertility has a strong medium-term positive effect on the growth of per capita income due to the supply of labor” [11]. According to Allen C. Kelley and Robert M. Schmidt, “the short-term costs (benefits) of fertility (lower mortality) significantly increased (decreased) in the 1980s, and the beneficial effects of past births on labor are not fully compensated” [12]. A study by

Jeffrey G. Williamson noted that “demographic dividends are not just the effect of the level of participation in the workforce, but also the growth effect” [13]. David E. Bloom, David Canning, Günther Fink, analyzing the different nature of the impact on the growth rate of the global problem of increasing the proportion of people aged 60 years and older in 2005-2050 for OECD and non-OECD countries, they conclude that the latter can mitigate economic consequences through behavioral reactions and political reforms [14].

The demographic situation in Russia is also difficult [15, p. 21]. In most regions of Russia, there is currently a decrease in the population, which leads to a decrease in the number of labor resources. Along with the decrease in the number of labor resources, negative changes are occurring in the specific gravity of the population aged younger and older than working age. The proportion of the population of young ages, that is, potential labor resources is declining, and the proportion of the population of retirement age is growing rapidly.

A decrease in the working-age population reduces the economic potential of the region and can lead to labor shortages. In many regions, there is a problem of reducing the quality of labor resources due to an increase in the incidence of the population at working age [16].

Lipetsk region belongs to demographically problematic regions. For the period 01.01.2000 - 01.01.2019 the population of the region decreased from 1233,7 thousand people to 1144 thousand people. The performed population forecast of the Lipetsk region by the method of moving the ages for the next two decades in three versions (medium, low, high) shows that the decrease in the total population of the region will continue at an increasing rate: according to the average forecast, from 0,6% in 2017-2022 up to 4,3% in 2032-3037 at the beginning of the period [17].

Since the population is both a source and a consumer of a regional public product, the difficult demographic situation in the region to a certain extent negatively affects its socio-economic development. There is a need for a detailed analysis of the impact of demography, labor and employment indicators on a regional social product in the development and implementation of strategic, long-term and short-term plans for the development of the socio-economic system of the region.

2. METHODOLOGY OF RESEARCH

In accordance with the basic principles of economic theory, the main factors affecting the volume of a social product (output) are labor costs, capital costs in the form of investments in the economy and scientific and technological progress. This process was thoroughly investigated and substantiated by such economists as Abramovich, Bombach and Solow [18]. The methodology for assessing the production potential of various levels continues to improve [19; 20].

The use of mathematical methods to analyze the social product (GDP - at the level of the economy as a whole, GRP - at the level of regions) allows a more detailed

approach to the problem of its formation, to trace possible changes in dynamics, to identify the relationship between components that have a direct impact on this complex indicator, get a quantitative assessment of the result and formulate scientifically based conclusions.

To achieve these goals, a correlation-regression analysis was chosen — the classical method of stochastic modeling of economic activity [21]. The legitimacy of using this approach is explained by the fact that the selected data can be considered random, and selected from the general population, distributed according to the multidimensional normal law. Using the capabilities of correlation and regression analysis allows us to determine both the combined effect of a combination of interrelated and simultaneously acting economic attributes - factors of various nature, and the individual effect of each of them on the formation of GDP (GRP). The result of this assessment is the regression equation, which describes the form of the relationship and allows you to obtain predicted values.

In our study, the base for conducting correlation and regression analysis is the GRP of the Lipetsk region.

The empirical basis of the study was the official statistical data of the Territorial Authority of the Federal State Statistics Service for the Lipetsk Region (then - Lipstat).

When determining the factors that influence the studied indicator, the approaches described in the articles by Kudymov V. [22], Magomadov N. S., Shamilev S. R. [23], Gusarova O. M., Kuzmenkova V. D., Komarov P. I. [24], Ivchenko Yu. S. [25].

The construction and evaluation of regression models was carried out in the systems STATISTICA and Microsoft Excel, using step-by-step algorithms for including significant variables.

As indicators influencing GRP (effective factor - y), five indicators-factors were selected, which act as factors-signs for GRP (input parameters):

x_1 – investments in fixed capital (in current prices, mln. rub.);

x_2 – cost of fixed capital (end of year; the full book value, mln. rub.);

x_3 – the number of companies and organizations (end of year, units);

x_4 – economically active population / labor force aged 15-72 years (according to labor force surveys, thousands of people);

x_5 – average annual number of employees (thousands of people).

The indicators were taken in dynamics for the period from 1999 to 2017.

Theoretical research methods used in this work included analysis, comparison, synthesis and generalization of scientific results, retrospective and statistical analysis.

3. RESULTS OF RESEARCH

The study included three stages.

3.1. Identification time trend model parameters

At the initial stage, trends were identified and trend equations were selected for each of the parameters (fig. 1).

For some of the indicators, several versions of the models were constructed that adequately reflect the dynamics of this factor.

The analysis of trend lines gives a clear idea of the direction and positive dynamics of each indicator.

The values of the determination coefficients in the constructed models are quite high and confirm that the time factor has a significant effect on the variation of each of the studied factor signs.

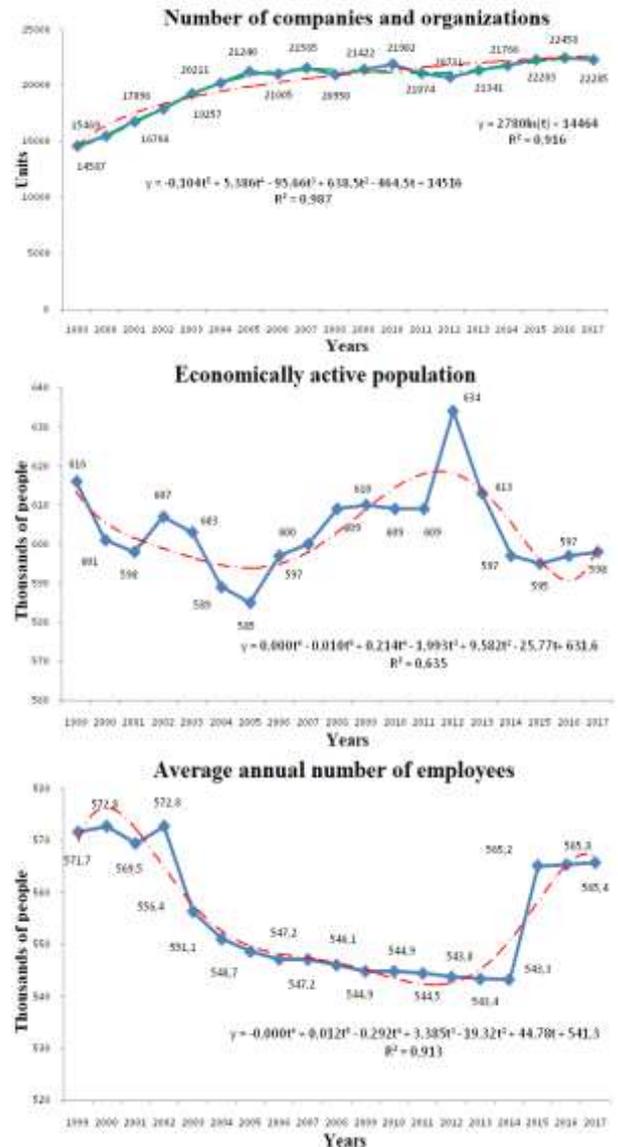
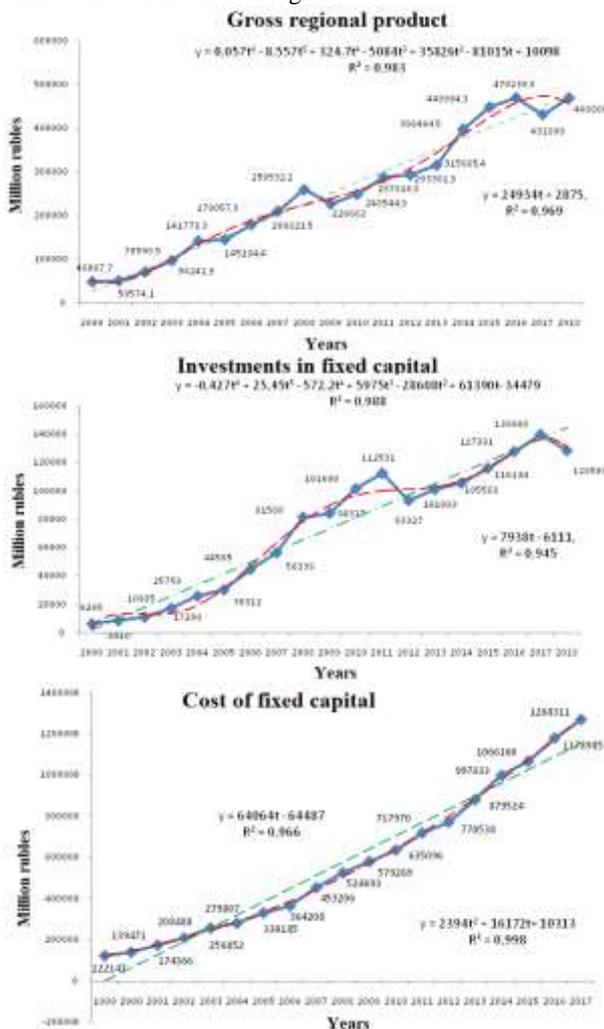


Figure 1 Identification of trends in model parameters according to Lipstat data, 1999-2017.

Note that the linear model describes well the dependence of the dynamics of GRP, investments in fixed capital, and the cost of fixed capital: the average rate of change in the level of GRP is 24934 million rubles, investments in fixed capital - 7938 million rubles, the cost of fixed capital - 64064 million rubles.

The revealed linear relationship indicates a uniform change in these indicators over time. Despite its simplicity, the linear model in practice is directly used to obtain predicted values and gives a completely satisfactory result.

The polynomial dependence adequately describes the dynamics of all the studied indicators, as we see from figure 1, such a model approximates the initial data series well, however, it does not always give a good result in forecasting, however, a polynomial of any order reduces to linear regression with its methods for estimating

parameters and testing hypotheses.

A semilogarithmic trend is typical for describing the trend in the number of enterprises and organizations and in this case describes a process of slowing growth ($a_1 = 2780 > 0$), in addition, the following economic interpretation is possible: when the time parameter t is increased by 1%, the number of enterprises and organizations grows by 27,8 units. The semilogarithmic model can be reduced to a linear form using the appropriate transformations.

3.2. Correlation analysis of model parameters

At the second stage, to identify the presence of multicollinearity of factor signs, the degree of influence of each indicator x_i ($i = 1, \dots, 5$) on the dependent variable y and to quantify the tightness of communication, a matrix of pair correlation coefficients was constructed (Table 1).

Table 1 Matrix of paired correlation coefficients

	Y	x_1	x_2	x_3	x_4	x_5
Y	1					
x_1	0,959	1				
x_2	0,989	0,953	1			
x_3	0,803	0,813	0,754	1		
x_4	-0,034	0,134	0,030	-0,167	1	
x_5	-0,255	-0,394	-0,214	-0,632	-0,184	1

The calculated pair correlation coefficients confirm a strong linear relationship between GRP and investments in fixed capital, the cost of fixed capital, the number of companies and organizations, and there is practically no correlation between GRP and the economically active population / labor force aged 15-72 years and the average annual number of employees.

3.3. Regression analysis of model parameters

At this stage, the classical model of multiple regressions was built, including all variables:

$$Y = 162739,15 + 0,71x_1 + 0,29x_2 + 5,96x_3 - 840,60x_4 + 441,15x_5.$$

Since the factors differ in their essence and have different units of measurement, for a more accurate assessment of the influence of factors on the result, a multiple regression model was constructed on a standard scale:

$$Y = 0,223x_1 + 0,714x_2 + 0,0953x_3 - 0,063x_4 + 0,035x_5.$$

The Fisher multiple regression model at the 5% significance level is statistically reliable ($F_{cr}(5,13) = 254,45$), i.e. regression coefficients are jointly significant (Table 2).

The statistical significance of the equation at a 5% significance level was checked using the coefficient of determination and the Fisher criterion ($F_{cr}(5,13) = 254,45$). It was established that in the studied situation 99,3% of the total variability of the factor-result – GRP is explained by the factors x_i ($i = 1, \dots, 5$). It was also established that, according to Student's test, at 5% significance level, the significance of the parameter is confirmed for the variable x_2 - the cost of fixed capital. The value of the average approximation error of 8,49% < 15% indicates a good quality of the selected model of the equation.

Table 2 Regression statistics for the 5 factor model

Multiple R	0,997
R-squared	0,993
Normalized R-squared	0,992
Standard error	145093,48
Average absolute approximation error	8,49

Taking into account the revealed multicollinearity and step-by-step algorithms for the inclusion of significant variables, a particular model was constructed including variables x_3, x_4, x_5 :

$$Y = -7299565,97 + 74,73x_3 + 3585,757x_4 + 6949,763x_5$$

and on a standard scale:

$$Y = 1,195x_3 + 0,267x_4 + 0,55x_5.$$

The significance of the private model as a whole (according to the Fisher criterion — $F_{cr}(5,13) = 254,45$ and the coefficient of determination) and its parameters (Student's test) is confirmed at a 5% significance level (Table 3).

Table 3 Regression statistics for the 3 factor model

Multiple R	0,901
R-squared	0,811
Normalized R-squared	0,773
Standard error	69832

4. DISCUSSION OF RESULTS

From the data of table 1, taking into account the Cheddock scale, we conclude that the relationship between:

- 1) GRP and factors x_1, x_2, x_3 strong;
- 2) GRP and factor x_5 reverse weak;
- 3) GRP and factor x_4 almost nonexistent.

The greatest influence on GRP is exerted by factor x_2 . In addition, the obtained values of pair correlation coefficients indicate that:

— there is multicollinearity of factors x_1, x_2 and x_3 ($|r| > 0,7$), in order to eliminate duplication, eliminate the effect of these indicators in unison and improve forecasting accuracy, it is advisable to include only one of them in the analysis;

— the size of the labor (factors x_4 and x_5) has little effect on GRP. This fact may indicate that GRP is provided with the quality of labor force training and working conditions (automation, digitalization).

Checking the significance of paired correlation coefficients using Student's criterion (at a 95 percent significance level) showed that:

1) $r_{yx1}, r_{yx2}, r_{yx3}$ – significant;

2) r_{yx4}, r_{yx5} – not significant.

For comparability of qualitatively dissimilar units of measurement of factor attributes, we commensurate indicators of the factor's relationship with the result were calculated — partial elasticity coefficients, which show how much on average the attribute-result of GRP changes on average with an increase in attribute-factor x_i ($i = 1, \dots, 5$) by 1% of its average level with a fixed position of other factors of the model. This approach made it possible to rank the factors by the force of influence on the result (Table 4).

Table 4 Partial coefficients of GRP elasticity by factors

Coefficient of elasticity	Ex_1	Ex_2	Ex_3	Ex_4	Ex_5
Value	0,203	0,708	0,518	-2,18	1,052

The calculation of the partial elasticity coefficients shows that the characteristic factors x_4 and x_5 separately significantly affect the GRP, since $|Ex_4|$ and $|Ex_5| > 1$.

According to the characteristic factors x_1, x_2, x_3 , the connection is inelastic.

We give an economic interpretation of the results:

— with an increase in x_1 - investments in fixed capital by 1%, the GRP of the Lipetsk region will increase by an average of 0,203% with the same other factors;

— with an increase in x_2 - the cost of fixed capital by 1%, the GRP of the Lipetsk region will increase by an average of 0,708% with the same other factors;

— with an increase in x_3 - the number of companies and organizations by 1%, the GRP of the Lipetsk region will increase by an average of 0,518%;

— with an increase in x_4 - of the economically active population / labor force aged 15-72 years by 1%, the GRP of the Lipetsk Region will decrease by 2,18% on average;

— with an increase in x_5 - the average annual number of employees by 1%, the GRP of the Lipetsk region will increase by 1,052% on average.

We interpret the economic significance of the parameters of the standardized classical model of multiple regression:

— $\beta_2 > \beta_i$ ($i=1,3,4,5$), then factor x_2 - the cost of fixed capital stronger than other factors of the system affects

the effective variable y - GRP of the Lipetsk region;

— the relationship between the result of y - GRP of the Lipetsk region and x_4 - of the economically active population / labor force aged 15-72 years is reversed;

— the relationship between the result y - GRP of the Lipetsk region and the factors x_1, x_2, x_3, x_5 is direct.

Based on the standardized coefficients and correlation coefficients, the coefficients of separate determination are calculated (Table 5).

Thus, due to variation x_1 - investments in fixed capital, 21,3% of the GRP variation in the Lipetsk region is explained, due to variation x_2 - the cost of fixed capital – 70,6%, due to variation x_3 - the number of companies and organizations – 7,7% , the variation of the remaining factors of the system accounts for less than 1%.

Table 5 Separate determination coefficients for the 5 factor model

Separate determination coefficient	d^2_{x1}	d^2_{x2}	d^2_{x3}	d^2_{x4}	d^2_{x5}
Value	0,213	0,706	0,077	0,002	-0,009

Using step-by-step algorithms for including significant variables, a standardized partial model of multiple regression was obtained:

$$Y = 1,195x_3 + 0,267x_4 + 0,55x_5.$$

We give an economic interpretation of the results:

— $\beta_3 > \beta_i$ ($i=4,5$), then factor x_3 - the number of companies and organizations stronger than other factors of the system affects the effective variable y - GRP of the Lipetsk region;

— the relationship between the result of y - GRP of the Lipetsk region and factors x_3, x_4, x_5 is direct.

Based on the standardized coefficients and correlation coefficients, the coefficients of separate determination are calculated (Table 6).

Table 6 Separate determination coefficients for the 3 factor model

Separate determination coefficient	d^2_{x3}	d^2_{x4}	d^2_{x5}
Value	0,96	-0,009	-0,14

Thus, due to the variation of x_3 - the number of enterprises, 96% of the variation in GRP of the Lipetsk region is explained.

5. CONCLUSIONS

According to the results of the study, causal relationships between the studied economic indicators using correlation

and regression analysis are revealed, which allow you to get quite extensive analytical information on the formation of GRP and indicators of socio-economic development of the Lipetsk region. Established the following.

Firstly, capital factors (especially investments in fixed capital and the value of fixed assets) have a stronger effect on the social product than labor factors.

Secondly, the factor x_4 (the number of economically active population) has the weakest connection with the resulting indicator. This can be explained either by indirect influence through the indicator of the number of people employed in the economy (which requires its exclusion from the model as a duplicate indicator), or by its incorrect reflection, which excludes accounting for qualitative characteristics (which requires its replacement or refinement through additional indicators).

Thirdly, the partial coefficients of GRP elasticity showed high values precisely for labor factors. Therefore, it remains important to study the problems of employment, demography and labor resources of the Lipetsk region in order to achieve an increase in GRP.

Prediction of regional economic systems using digital technologies, the use of mathematical modeling allows you to make managerial decisions on a scientific basis. The proposed science-based model can be used as the basis for the development of strategic plans and programs of the Lipetsk region.

The choice of indicators is not final, and can be interpreted depending on the level of socio-economic development of the regions.

ACKNOWLEDGMENT

Study was supported by Russian Foundation for Basic Research, project №18-410-480003 p_a "The impact of demographic and migration processes on the regional employment system and gross regional product"

REFERENCES

- [1] Stefanova N.A., Sedova A.P., Model of the Digital Economy, Karelian Scientific Journal, 2017, vol. 6, No. 1 (18), pp. 91–93.
- [2] Kozyrev A.N., Mathematical models and forecasting in the digital economy, Access Mode: http://spkurdyumov.ru/digital_economy
- [3] Lawrence R.Klein, The statistical approach to economics, Journal of Econometrics, Volume 37, Issue 1, January 1988, Pages 7-26. Available at SSRN: [https://doi.org/10.1016/0304-4076\(88\)90071-1](https://doi.org/10.1016/0304-4076(88)90071-1).
- [4] Clopper, A. (2012), The Craft of Economic Modelling. Fifth edition. Interindustry Economic Research Fund, Inc. p. 149. Available at SSRN: <http://inforumweb.umd.edu/papers/publishedwork/books/craft1.pdf>.
- [5] Veduta E.N., Dzhakubova T.N., Economic models in the digital economy, Public Administration in the 21st Century: Compendium, Electronic Edition of Network Distribution / Team of Authors. - M.: "KDU", "Dobrosvet", 2018. Access mode: https://bookonline.ru/system/files/node/lecture/1-14_veduta_86-91.pdf
- [6] Titova O.V., Kozlova E.I., GRP as a tool for analyzing the structural proportions of a regional economy, Modern economy: problems and solutions, 2016, No8 (80), S. 104–113.
- [7] Mikhailov V.V., Sargsyan V.V., Analysis of the multidimensional factor model of GRP and industry models of the dependence of GRP and financial results on investment and production activity, Siberian financial school, 2003, No. 4 (49), pp. 15-21.
- [8] Kozlov A.A. Socio-demographic factors of regional development, Vestnik TSU, vol. 12, no. 6, 2007. pp. 669–671.
- [9] Morozova E.A., Chelombitko A.N., Andreeva L.M., Demographic situation and its impact on the socio-economic development of the region, Bulletin of the Kemerovo State University, 2012, No2 (50), pp. 213-219.
- [10] David E. Bloom & Richard B. Freeman. Economic development and the timing and components of population growth. Journal of Policy Modeling. Volume 10, Issue 1, April 1988, Pages 57-81.
- [11] James A. Brander & Steve Dowrick. The Role of Fertility and Population in Economic Growth: Empirical Results from Aggregate Cross-National Data. Journal of Population Economics Vol. 7, No. 1 (Feb., 1994), pp. 1-25
- [12] Kelleyand A.&Schmidt R. Aggregate population and economic growth correlations: The role of the components of demographic change. Demography, 1995, vol. 32, issue 4, 543-555.
- [13] Williamson, Jeffrey G., Demographic Dividends Revisited (August 22, 2013). Asian Development Review 30:2. Available at SSRN: <https://ssrn.com/abstract=2758889>.
- [14] David E. Bloom, David Canning, Günther Fink. Implications of population aging for economic growth.

National bureau of economic research 1050 Massachusetts Avenue Cambridge, MA 02138 January 2011. Available at SSRN: <https://www.nber.org/papers/w16705>.

[15] Elin A.M., Pashin N.P., Problems of demography and ways to solve them in modern Russia, *Bulletin of science and education*, 2019, No 17 (71), pp. 19-28.

[16] Vishnevsky A., Mortality in Russia: the failed second revolution, *Demographic Review*, Electronic scientific journal, 2014, Volume 1, No. 4. pp. 5–40. Access mode: <https://demreview.hse.ru/issue/view/154>.

[17] Novak M. A., Kozlova E. I. Application of the Method of Moving Ages to Predict the Population of the Lipetsk Region // 2019 1st International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA). Lipetsk, Russia, 2019, pp. 224–228.

[18] Solow R. Technical change and the aggregate production function // *Review of Economics and Statistics*. 1957. №39 (3).

[19] The Production Function Methodology for Calculating Potential Growth Rates & Output Gaps / K. Havik, K. McMorrow, F. Orlandi, Ch. Planas, R. Raciborski, W. Röger, A. Rossi, A. Thum-Thysen, V. Vandermeulen. *European Economy. Economic Papers* 535. November 2014. Brussels. PDF. 120pp. Available at SSRN: http://ec.europa.eu/economy_finance/publications/economic_paper/2014/pdf/ecp535_en.pdf].

[20] E.-A. Andrei, E. Bugudui. *Econometric Modeling of GDP Time Series. Theoretical and Applied Economics*. Volume XVIII (2011), No. 10(563), pp. 91-98. Available at SSRN: <http://store.ectap.ro/articole/652.pdf>.

[21] *Econometrics: a textbook for masters* / I.I. Eliseev [and others]; under the editorship of I.I. Eliseeva, M. : Publishing house Yurayt, 2014, 453 p.

[22] Kudymov V. M., Interrelation of socio-economic processes with an indicator of gross regional product, *Regional economics - theory and practice*, 2007, No.15, pp. 42-51.

[23] Magomadov N. S., Shamilev S. R., Analysis of the dynamics of GRP of the regions of the Russian Federation by production functions, *Modern problems of science and education*, Electronic scientific journal, 2014, No.6. Access mode: <https://www.scienceeducation.ru/pdf/2014/6/194.pdf>.

[24] Gusarova O.M., Kuzmenkova V.D., Komarov P.I. Digital models of socio-economic development of regional subjects, *Fundamental Research*, 2018, No. 8, pp. 42–47.

[25] Ivchenko Yu. S., Determination of the main factors of the gross regional product level by econometric modeling methods for the aggregate of regions of the Russian Federation, *Statistics and economics*, 2019, Vol. 16, No. 6, pp. 4-18.