European Proceedings of
Social and Behavioural Sciences
EpSBS

www.europeanproceedings.com
e-ISSN: 2357-1330

DOI: 10.15405/epsbs.2020.04.22

PEDTR 2019

18th International Scientific Conference "Problems of Enterprise Development: Theory and Practice"

POPULATION REPRODUCTION INDICATORS' REGRESSION MODELS IN THE RUSSIAN FEDERATION

E. Y. Nuykina (a)*
*Corresponding author

(a) Samara State University of Economics, 443090, Soviet Army Str., 141, Samara, Russia, nuikina1973@mail.ru

Abstract

Demographic processes on the Russia territory are based on huge differences between the country subjects according to natural, climatic, economic, social, national, political and other conditions. Demographic processes features such that the relationship between demographic phenomena and the influence exerted on them by other phenomena can be studied only by the theory means, which allows to detect and quantify these relationships and communications. The system of vital population movement quantitative and qualitative indicators characterizes the country's demographic situation, in the region, i.e. the state of population structures demographic processes, which determine the conditions for its further reproduction. The purpose of this study is to identify the factors role determining the Russian Federation demographic situation territorial differences on the multi-factor regression models construction basis. Indicators such as life expectancy (number of years), demographic load ratio (persons under working age per 1000 persons of working age), natural growth (per 1000 population) are considered as the object of population reproduction modelling. The factors selection to construct a particular correlation-regression model is based on a detailed analysis of the paired correlation coefficients matrix between the dependent variable and all the argument factors considered, and the collinear factors identification. The practical implementation of multidimensional statistical analysis was carried out using the MS Excel package. Regression models can serve to develop demographic projections in the future. The work used data of the Russian Federation Federal State Statistics Service.

Keywords: Multiple-factor regression model, demographic situation, population' reproduction.

https://doi.org/10.15405/epsbs.2020.04.22
Corresponding Author: E. Y. Nuykina
Selection and peer-review under responsibility of the Organizing Committee of the conference
eISSN: 2357-1330

1. Introduction

The regional economy and social sphere phenomena and processes become sufficiently aware only if, together with their essence meaningful analysis, it is possible to give a quantitative expression of their inherent objective patterns and interrelationships, explaining the reasons for their deep territorial differentiation (Maleva, Makarentseva, & Tretyakova, 2017). A prerequisite for obtaining opportunities to control and predict the population reproduction process in the Russia regions, to scientifically justify the choice of regional policy related to these processes, is the development of the relevant indicators statistical model's set (Kurdova & Efimova, 2018).

One of the most important aspects of the mathematical and statistical apparatus application in the field of social phenomena under consideration is the relationships existence in the statistical patterns form. The latter manifest themselves in mass social processes, in a social objects combination consisting of a units large number, and are formed according to the large numbers law. The combined influence of multiple random factors and causes leads to results that are virtually case-independent. Statistical patterns express causal and other relationships inherent in this field of social phenomena and processes.

2. Problem Statement

Demographic processes features such that the relationship between demographic phenomena and the influence exerted on them by other phenomena can be studied only by the theory means, which allows to detect and quantify these relationships and communications. (Ivliev & Cheremisina, 2014). Indicators such as life expectancy (number of years), demographic load ratio (persons under working age per 1000 persons of working age), natural growth (per 1000 population) are considered as the object of the Russian Federation population reproduction modelling.

3. Research Questions

The modeling object specificity is determined not only by the dependent quantities composition, but also by the factors set that determine them. In demographics, one can encounter such cases when the answer to the question of which factor is the cause and which consequence is complicated by the circumstances. Most often in these cases we deal with factors interdependence, that is, cases where factor signs affect the effective, and those in turn have the opposite effect.

4. Purpose of the Study

The purpose of this study is to identify the factors role determining the Russian Federation demographic situation territorial differences on the multi-factor regression models construction basis. The study of population reproduction factor links is carried out according to one of the most priority, in our opinion, indicators of natural population movement, such as life expectancy, demographic load ratio, natural growth.

https://doi.org/10.15405/epsbs.2020.04.22
Corresponding Author: E. Y. Nuykina
Selection and peer-review under responsibility of the Organizing Committee of the conference
eISSN: 2357-1330

5. Research Methods

Two directions should be highlighted as part of the methods for constructing multi-factor regression models. The first approach algorithm consists in a multi-step analysis, on each iteration of which a new, additional factor-argument is included in the model, the significance of each regression coefficient according to the Student's t-test and the sign convergence at each coefficient of the multiple regression equation with the paired correlation coefficient sign are checked. The model construction begins with a factor-argument consideration most closely related to the dependent variable, and ends with the condition that the resulting model is adequate (the calculated value of the F-criterion becomes larger than the table value) (Ragsdale & Gutenkunst, 2017).

Another approach is to analyze in detail the matrix of paired correlation coefficients between the dependent change and all argument factors considered, and to identify collinear (interdependent) factors, one of which should be excluded from the factors composition (Garthoff & Otto, 2018). The idea of multi-step analysis is to carry out a calculated iterations number, during which all factors selected by theoretical analysis are sequentially included in the model and measures of materiality of influence on the all factors effective indicator taken into account by the model and results convergence are evaluated at each calculations step (Roth, & Caswell, 2016). Factors that are not essential and do not contribute to improving the performance characteristic calculated and actual values convergence are omitted and replaced by new ones (Kobzar, 2006).

This approach was chosen for factor link analysis. The modeling object specificity is determined not only by the dependent quantities composition, but also by the set of factors that determine them. In demographics, it is possible to meet cases where the answer to the question of which factor is the cause and which consequence is difficult by the incoming circumstances. Most often in these cases we deal with the interdependence of factors, that is, cases where factor signs affect the effective, and those in turn have the opposite effect (Wilson, 2016). The factors mutual effect can be seen in this example: women 's employment is an obstacle to motherhood and motherhood is an obstacle to employment.

As a result of a wide range of factor characteristics logical consideration, the following argument factors composition is obtained, which are most closely related to simulated values:

X1 - men and women ratio (per 1000 men women);

X2 - urban population proportion in the total population (%);

X3 - abortion (number of abortions per 100 births);

X4 - marriages and divorces ratio (1000 marriages account for divorces);

X5 - proportion of the population older than working age (%);

X6 - population provision with doctors (number of doctors per 10000 population);

X7 - housing provision (m^2 area per inhabitant);

X8 - preschool institutions number;

X9 - average per capita income (rubles);

X10 - proportion of the population with cash incomes below the subsistence level in the total population (%);

X11 - general unemployment level (in % of the total economically active population);

https://doi.org/10.15405/epsbs.2020.04.22
Corresponding Author: E. Y. Nuykina
Selection and peer-review under responsibility of the Organizing Committee of the conference
eISSN: 2357-1330

X12 - population incidence (registered patients with a diagnosis established for the first time in their lives, per 1000 population);

X13 - recorded murders and attempted murders number;

X14 - paved public roads density (km of roads per 1000 km^2 of territory);

X15 - consumption of meat and meat products per capita (kg);

X16 - grain yield (t / ha);

X17 - GRP per capita (thousand rubles);

X18 - average annual air temperature, ^0C.

As a result of multi-step regression analysis, the equation of territorial life expectancy levels multiple regression is characterized by the expression:

$У_1$=62,58-0,0717X_{11}-0,2014X_{12}+0,0049X_{14}+0,0855X_{16}+0,1057X_{18},

The population's age composition has a great influence on the population reproduction rates. Boyarsky (1975) believed that the age composition is in some way a vital record. The constructed demographic load coefficient model has a high multiple correlation coefficient R = 0,909. The calculated Fisher criterion is 40,9.

$У_2$=175,32+0,8715X_2-0,127X_3-0,118X_4+0,808X_7-0,026X_{10}+0,0496X_{11}+0,0191X_{12}

The natural population growth rate depends on fertility and mortality, as well as the age structure:

$У_3$=35,92-0,011X_1-0,034X_2-0,026X_3-0,378X_5-0,797X_7++0,013X_{11}

Private elasticity factors are also used to compare the role of different factors in the territorial differences formation in the demographic situation. They show how many percent the dependent variable changes on average with the factor characteristic x_i changing by 1% with the other factors unchanged. Calculation is made according to the following formula:

$$E = b_i * \frac{\overline{X_i}}{\overline{Y}},$$

where: b_i - the regression equation coefficient;

$\overline{X_i}$ - average value of *i*-factor characteristic;

$\overline{Y}$ - average value of effective characteristic.

Table 01 gives the calculated values of the partial elasticity coefficients of life expectancy (number of years) У1, demographic load ratio (persons under working age per 1000 persons of working age) У2, natural growth (per 1000 population) У3 considered models.

Table 01. Private elasticity factors of vital indicators and their determinants

Factors arguments	Llife expectancy $У_1$	Demographic load ratio $У_2$	Natural growth $У_3$
X_1	-	-	-0,715
X_2	-	0,178	-0,404
X_3	-	-0,06	-0,81
X_4	-	-0,163	-
X_5	-	-0,416	-1,29

https://doi.org/10.15405/epsbs.2020.04.22
Corresponding Author: E. Y. Nuykina
Selection and peer-review under responsibility of the Organizing Committee of the conference
eISSN: 2357-1330

X_6	-	-	-1,27
X_7	-	-0,14	-
X_8	-	-	-
X_9	-	-	-
X_{10}	-	-0,03	-
X_{11}	-0,1	0,02	0,04
X_{12}	-0,2	0,04	-
X_{13}	-	-	-
X_{14}	0,199	-	-
X_{15}	0,186	-	-
X_{16}	-	-	-
X_{17}	-	-	-
X_{18}	0,09	-	-

Source: author based on (Federal State Statistics Service, 2019)

According to the specific elasticity factors (Table 01), the priority of the models argument factors under consideration is thus formed:

Model 4: E12, E14, E16, E11, E18.

Model 5: E5, E2, E4, E7, E3, E12, E10, E11.

Model 6: E5, E7, E3, E1, E2, E11.

6. Findings

The multiple regression parameters equation of the population territorial life expectancy levels mean that with an increase of the general unemployment rate by 1%, life expectancy decreases by 0,0717 years; life expectancy decreases by 1 years with population morbidity per 0,2014 unit. The paved roads density has an impact on improving life expectancy. Among the life expectancy argument factors model are natural and climatic factors: grain yield (X16), average annual air temperature (X18), although their impact is not so significant (dy4x16=0,105; dy4x18=0,094). The model is adequate by F-criterion, as its calculated value (10,06) exceeds the table value. The multiple correlation coefficient is 0,856. The determination coefficient shows that the territorial variation of the characteristic under consideration by 73,4% is determined by the factor characteristics variation included in the model. Life expectancy is the most elastic relative to the population incidence rate.

Population urbanization (X2), expressed by the proportion of the urban population in the total population (direct dependence), has the greatest impact on the demographic burden territorial differentiation. With the increase of this indicator by 1 p.p. the demographic load factor (persons younger than working age per 1000 people of working age) increases by 0,8715 units. This effective characteristic is also influenced by the factor-argument of the population housing availability (m2 of area per resident). Increase of this index by 1 m2. Increases factor characteristic by 0,8 units.

The higher the proportion of young ages in the total population, the more births, all other things being equal. Each of the indicators listed is in itself effective, so by performing a multiple correlation analysis on the factor-arguments we have proposed, it can be concluded that the natural population growth indicator has a very high elasticity with respect to many factor-arguments under consideration. Higher or

lower regional levels of population over working age (X5), housing provision (X7), abortion (X3), men and women ratio (X1) by 1% leads to a decrease in the natural population growth by 1, 29; 1.27; 0,813 and 0,715 %, respectively.

7. Conclusion

Demographic factors are very diverse in their content and can manifest themselves in the economic, political, cultural, socio-psychological spheres of public life (Hallett, Farrer, Suding, Mooney, & Hobbs, 2018). In general, correlation regression analysis shows that many factors affecting vital indicators are constantly affecting. Socio-demographic factors of the Russian Federation regions play the most prominent role in territorial differentiation formation of the population natural movement. The family institution crisis in modern society is reflected in the demographic situation sharp deterioration. Socio-economic factors are the next most important factors. Insufficient housing provision combined with low incomes of the region's inhabitants majority has a negative impact on the population reproduction nature. The demographic situation is a phenomenon secondary to the objective patterns governing population development in the given socio-economic conditions. It is shaped by a number of long-term trends in population development and represents a stage in population development (Fernihough, 2017). Regression models can be used to develop population projections, the need for which is linked to the tasks of forecasting and planning socio-economic processes in general.

References

Boyarsky, A. Y. (1975). Population and methods of its study. Retrieved from http://www.demoscope.ru /weekly/knigi/polka/gold_fund011.html. Accessed: 11.12.2019. [in Rus.].

Federal State Statistics Service (2019). Demographics. Retrieved from: https://www.gks.ru/folder/12781. Accessed: 29.09.2019. [in Rus.].

Fernihough, A. (2017). Human capital and the quantity – Quality trade-off during the demographic transition. *Journal of Economic Growth, 22*, 35-65. DOI:10.1007/s10887-016-9138-3

Garthoff, R., & Otto, P. (2018). Verfahren zur Überwachung räumlicher autoregressiver Prozesse mit externen Regressoren [Statistical surveillance of spatial autoregressive processes with exogenous regressors]. *AStA Wirtschafts- Und Sozialstatistisches Archiv, 12*(2), 107-133. DOI:10.1007/s11943-018-0224-1

Hallett, L. M., Farrer, E. C., Suding, K. N., Mooney, H. A., & Hobbs, R. J. (2018). Tradeoffs in demographic mechanisms underlie differences in species abundance and stability. *Nature Communications, 9*, 5047. DOI: 10.1038/s41467-018-07535-w

Ivliev, M. I., & Cheremisina, N. V. (2014). Demographic situation in modern Russia. *Socio-Economic Phenomena and Processes, 9*(8), 48-53. [in Rus.].

Kobzar, A. I. (2006). *Applied mathematical statistics. For engineers and scientists.* Moscow: Fizmatlit. [in Rus.].

Kurdova, M. A., & Efimova, O. G. (2018). Modern demographic policy in Russia and ways of its implementation. *Economics and Business: Theory and Practice, 2,* 51-55. [in Rus.].

Maleva, T. M., Makarentseva, A. O., & Tretyakova, E. A. (2017). Pronatalist demographic policy in the eyes of the population: Ten years later. *Ekonomicheskaya Politika, 12*(6), 124-147. DOI:10.18288/1994-5124-2017-6-06 [in Rus.].

Ragsdale, A. P., & Gutenkunst, R. N. (2017). Inferring demographic history using two-locus statistics. *Genetics, 206*(2), 1037-1048. DOI:10.1534/genetics.117.201251

Roth, G., & Caswell, H. (2016). Hyperstate matrix models: Extending demographic state spaces to higher dimensions. *Methods in Ecology and Evolution, 7*(12), 1438-1450. DOI:10.1111/2041-210X.12622

Wilson, T. (2016). Visualising the demographic factors which shape population age structure. *Demographic Research, 35,* 867-890. DOI:10.4054/DemRes.2016.35.29