DOI: 10.15838/esc.2021.3.75.14
UDC 314.37:332, LBC 60.7

Cluster Analysis of the Dynamics of the Birth Rate of Fourth and Subsequent Children in Russian Regions*

Svetlana N. KOSTINA
Ural Federal University named after the First President of Russia B.N. Yeltsin
Yekaterinburg, Russian Federation
e-mail: s.n.kostina@urfu.ru
ORCID: 0000-0001-9346-600X

Aleksandr V. TRYNOV
Institute of Economics, Ural Branch of the Russian Academy of Sciences
Yekaterinburg, Russian Federation
e-mail: Trynovv@mail.ru
ORCID: 0000-0002-2969-2536; ResearcherID: C-5565-2014

Abstract. The article presents results of the research in the differentiation of Russian subjects by birth rate of fourth and subsequent children, depending on the region's socio-economic development level for the period of 2005–2017. The authors conduct cluster analysis based on Ward's method of time series of 3 groups of indicators, particularly demographic, economic and social, in the regional context. As a result, they identify 6 clusters, describing the situation of Russian subjects, depending on the values of the birth rate of fourth and subsequent children and indicators of socio-economic development of the region for 2005–2017. The study reveals that in the period under review there are two main trends: first, transition of RF subjects from Cluster 1, characterized by sufficiently high values of the birth rate of fourth and

* The research is financially supported by the RFBR in the framework of the scientific project No. 19-010-00373 "Security and Pseudo-Security of the Economic Development of Russian Regions under the Influence of Multidirectional Trend".

For citation: Kostina S.N., Trynov A.V. Cluster analysis of the dynamics of the birth rate of fourth and subsequent children in Russian regions. *Economic and Social Changes: Facts, Trends, Forecast*, 2021, vol. 14, no. 3, pp. 232–245. DOI: 10.15838/esc.2021.3.75.14

subsequent children and low indicators of socio-economic development, to the clusters with a lower value of the coefficient and higher indicators of socio-economic development; second, transition of RF subjects from the clusters with a low birth rate of fourth and subsequent children to the clusters characterized by the increased birth rate against the background of improved socio-economic development. In the current period, it is possible to distinguish the formation of two poles of large families – these are "the poor with many children" regions, in which a high birth rate of fourth and subsequent children is associated with low socio-economic development, and "the rich with many children" regions with high birth rates and a high level of socio-economic development. Between them there are other RF subjects, which are gradually moving away from the pole of "the poor with many children", but have not approached the pole of "the rich with many children" yet. The novelty of the study lies in the application of the author's approach to the analysis of relationship between large families and socio-economic development of Russian regions in the temporal dimension.

Key words: birth rate, crude birth rate, birth rate of fourth and subsequent children, cluster analysis, having many children, socioi economic situation, subjects of the Russian Federation.

Introduction

In recent decades, the birth rate problem in Russia is critical and attracts attention of a wide range of specialists. Despite the 15-year active pro-natalist policy, the crude birth rate (CBR) has been declining in the Russian Federation since 2016[1]. There remains regional differentiation of the birth rate: the CBR value varies from 1.12 in the Leningrad Oblast to 2.97 in the Tyva Republic (according to the 2018 data)[2]. For the constituent entities and the country as a whole, the CBR differs in birth order significantly [1]. Though the number of births of first and second children is decreasing, the birth rate of third and subsequent children by birth order is growing quite steadily in relation to the previous period; the number of large families and the number of children in them are also going up [1].

The influence of regional socio-economic differences on the birth rate of first and subsequent children (including multi-child parenting) is quite complex and contradictory. The results of econometric modeling based on the dynamic series of panel data for RF subjects show that the significance of individual factors impact varies both for birth order coefficients and individual coefficients in dynamics [1]. The study is aimed at trying to cluster RF subjects by birth rate of fourth and subsequent children, depending on the socio-economic conditions in the period of 2005–2017. Its scientific novelty consists in the elaborated comprehensive approach to the analysis of fertility processes: first, the birth of fourth and subsequent children is a study object, second, the relationship between the birth of fourth and subsequent children and the socio-economic state of regions is considered, and third, the dynamics of its changes is evaluated.

Methodological approaches to clustering in the studies of socio-demographic processes

First, it should be noted that clustering birth rate-related processes is based on analyzing the impact, which various socially and historically determined factors have on them. Studies devoted to this problem are quite numerous [2], but their results are contradictory [1]. When determining a birth rate, the influence of various factors, such as economic, demographic, social, socio-cultural,

[1] Demographic results of the first half of 2019 in Russia (part I). Available at: http://www.demoscope.ru/weekly/2019/0823/barom05.php (accessed February 12, 2020).

[2] *Ibidem.*

and socio-psychological, is considered. They can be divided into macro-and micro- level factors. The macro-level factors are the following: GDP per capita, consumer confidence index [3], quality of life indicators [4], average per capita monetary income of the population, average monthly salary, per capita fixed capital investment, number of doctors of all specialties, average provision of housing for the population [5], number of women of reproductive age, etc. Micro-factors include characteristics of household, such income level [6], parents' education level, presence of a partner, availability of work for parents [7] , etc.

The analysis of regional differentiation is based on the ideas of spatial demography. In the modern period, its methodology has significantly expanded due to the use of a number of spatial analytics methods: spatial econometrics, geographically weighted regression, multi-level modeling, and spatial structure analysis. For example, the GWR (Geographically Weighted Regression) method allows us to study dynamic models of spatial and temporal clustering.

Spatial analysis is used to study differences in the course of demographic processes both between countries and between regions of one state. For instance, I. Mietule, I. Maksymova, and K. Holikova, having made cluster analysis, described the impact of dynamic changes inherent in the socio-demographic and economic characteristics of some European countries on business and the social sphere in individual clusters, using the example of Latvia and Ukraine [8].

A significant number of works are devoted to the intra-country spatial differentiation of socio-demographic processes. For example, P.J. Boyle, E. Graham, and Z. Feng singled out local birth rate clusters based on factors, such as place of residence, housing market characteristics, and social interaction processes [9].

The researchers I. Salvini, G. Gabrielli, and A. Paterno analyzed the relationship between mortality and birth rates and socio-economic characteristics, using dynamic factor analysis and cluster analysis methods, to describe convergence processes and heterogeneity among the clusters [10]. Cluster analysis was also used, for example, to work out the active demographic policy for the counties of Croatia [11].

Within the framework of spatial demography there appears a temporal dimension, which characterizes the dynamics of changes in the studied indicators in different territories. J. Pénzes, Z.I. Pásztor, and P. Tátrai analyzed databases of the 1980–2011 censuses and revealed spatial imbalances in the demographic characteristics of peripheral settlements and correlations between regional backwardness and population decline [12].

M. Yüceşahin and A. Tulga considered demographic changes and spatial clustering in terms of a qualitative approach (cluster analysis) in the Middle East and North Africa on the basis of selected demographic indicators for 1950, 1980 and 2015. One of the conclusions of this study is that MENA countries represent three different structures in different time periods in terms of demography, and these structures are responsible for regional social, economic, and political transformations [13].

In recent years, cluster analysis has also been widely used in Russian socio-demographic studies. A number of works are devoted to the cluster analysis of Russian regions, where the study objects are as such: level of human capital development [14], social potential of the region [15], socio-demographic situation (GRP, life expectancy, migration, morbidity and unemployment[3]), demographic behavior [16], demographic potential [17], birth and mortality rates of the population [18], functioning of healthcare systems [19] , etc.

[3] Gladyshev A.V. *Statistical analysis of the socio-demographic situation in the Siberian Federal District: Candidate of Economic Sciences Dissertation Abstract.* Moscow, 2005. 22 p.

Demographers, when applying the cluster approach, deal with at least two main methodological issues – identification of indicators and selection of the clustering method.

The variety of clustering methods used in demographic studies is significant. Clustering algorithms are divided into two categories: hierarchical and non-hierarchical. The results of hierarchical algorithms are presented in the form of a dendrogram – a tree diagram that shows in what order the objects are divided into clusters. Hierarchical algorithms, in turn, make up two groups: agglomerative and divisive [20]. Cumbersome nature of calculations is a disadvantage of hierarchical procedures.

In modern economic and demographic studies researchers often use the K-means method to obtain unbiased estimates for a large number of observation units [21] and the EM clustering algorithm to build data models using a linear combination of multidimensional normal distributions [15]. Popular geodemographic analysis algorithms include fuzzy geographically weighted clustering (FGWC) and its modification based on the ABC optimization scheme (FGWC-ABC) [22]. The fuzzy clustering algorithm has been recently gaining popularity, as well as a number of its modifications (fuzzy geographically weighted clustering, intuitionistic fuzzy geographically weighted clustering (MIPFGWC), Kernel Fuzzy Geographically Clustering (KFGC)) [23]. For example, fuzzy clusterization was used in the study of analysis of demographic potential of Russian regions [24] and effectiveness of regional healthcare systems [25].

Materials and methods

To analyze regional convergence of large families, we conducted clustering of RF subjects, based on the assessment of territorial differentiation of the indicator of the birth rate of fourth and subsequent children and the level of socio-economic development of the regions from 2005 to 2017. The choice of the basic variable was determined by a number of prerequisites. First, it is the "diversity" of modern ideas about the content of multi-child parenting: if classically a large family meant having five or more children, at present, within the framework of the Russian state's social policy, the boundary has shifted towards families with fewer children (three or more). Second, the impact of state support measures to encourage the birth of third children (implemented first at the regional, then at the federal level) on the birth rate has not been fully studied. Accordingly, the birth rate of fourth and subsequent children was chosen for the study, as it, on the one hand, characterizes a large number of children and, on the other hand, is to a lesser extent influenced by the state policy.

For clustering, we used indicators that can be divided into three groups. The choice of variables was determined, first, by the need to take into account various groups of factors, such as demographic, economic and social; second, by the availability of time series of statistical data for the selected indicators. As a result, the following indicators were used as regressors (explanatory variables):

1) economic – real per capita monetary income according to regional PPP, rubles per month, nominal GRP per capita, rubles per year, unemployment rate, share of the population with income below subsistence minimum, R/P 10% ratio (social stratification degree);

2) demographic – marriage rate, life expectancy at birth, migration growth, urban population share, divorce rate, number of abortions per 100 births;

3) social (infrastructure) – total area of residential premises per inhabitant, square meters, population per hospital bed, coverage of children with preschool education, Human Development Index.

To solve the tasks, we used official statistics data for Russian regions for 2005–2017. At the beginning of the study period, 45 regions were included in the sample. This is due to the lack of statistical data on the birth of fourth and subsequent children in other RF subjects. Then, as more data became available, the sample was replenished, and as of 2017 it comprised 85 regions of the Russian Federation. Besides, in the study we consider that the Arkhangelsk Oblasts includes Nenets Autonomous Okrug, and the Tyumen Oblast includes Khanty-Mansi and Yamalo-Nenets autonomous okrugs.

Still it is rather controversial to include regions in the sample, when the necessary statistical information on them becomes available. The analysis of scientific literature shows that this issue was practically not considered in previous works. The authors avoid this problem by truncating the sample by number of regions and by time. For example, in the work of O.M. Shubat and I.V. Shmarova [24], the database of 77 RF subjects for 2012–2015 was used. A.M. Ilyshev and A.P. Bagirov [25] conducted clustering of regions by level and factors of reproductive activity based on the 2006 data alone.

It should be noted that the issue of cluster dynamics (the transition of regions of the original cluster to another and the process of formation and identification of new clusters) is practically not studied in scientific literature. As a rule, clusters are represented by a static group of regions that are not subject to changes. For example, in the work of N.P. Tikhomirov and T.M. Tikhomirova [26], the regions were grouped into three homogeneous clusters, the composition of which did not change in the study period (2000–2019). In addition, several subjects were classified as "non-typical regions". Unfortunately, the authors did not provide a detailed explanation of this measure. In our opinion, the issue of cluster dynamics makes it possible to more accurately differentiate regional demographic policy measures. The fact of the transition of a region from one cluster to another, as well as the analysis of the factors ensuring such a transformation, is of considerable scientific interest and may be the object of further research. To solve this methodological problem, we conducted cluster analysis in two variants. The first variant included calculations for 45 regions in the dynamics from 2005 to 2017. In the second variant, the calculations were carried out for 85 RF subjects, as they were included in the sample based on the availability of data.

Statistical processing was performed using hierarchical cluster analysis according to Ward's method in Stata 15 application software package. This approach to clustering was first proposed by J.H. Ward [27]. The attractiveness of hierarchical cluster analysis for our study is due to the program's capability to independently determine the number of clusters to break down. The difference between Ward's method and other cluster analysis methods is that methods of variance analysis are used to estimate the distances between clusters. The distance between clusters is taken as the increase in the sum of squares of the distances of objects to the cluster center, obtained as a result of their union. That is, there is a merger of clusters, the union of which gives the smallest increase in intra-cluster dispersion, i.e., to the greatest extent leads to "loosening" of the clusters identified in the previous steps of the procedure. Thus, the clustering (joining) algorithm according to Ward's method includes four stages.

1. An average value of individual variables for objects included in the cluster is calculated.

2. Squared Euclidean distances between individual observations of each cluster and the average cluster value are calculated.

3. Received values are summed up.

4. Observations are grouped into clusters in such a way that the reduction in the total sum of the distances between the clusters is the smallest.

Results and discussions

In the course of calculations based on the sample of 45 regions, for which there are data for all years of the study period, 5 main clusters were identified *(Table 1)*. In the calculations involving 85 regions of the Russian Federation, 6 clusters were identified *(Table 2)*. The final distribution of regions by cluster is presented in *Table 3*. The selected clusters show the relationship between the birth rate of fourth and subsequent children and the socio-economic situation of the regions that changes in the period under review. Thus, the transitions of regions from one cluster to another are tracked over time.

Cluster 1. It is characterized by a relatively high birth rate of fourth and subsequent children and the lowest indicators of socio-economic development. In particular, there is the lowest value of average per capita income and GRP per capita, while the maximum average value of the poverty and unemployment level. It also has the smallest share of urban population. The number of regions in Cluster 1 gradually decreased from 40 to 3 in 2005–2013. Moreover, the regions not only moved from less "prosperous" clusters to more "prosperous" ones, but there was also a reverse movement – 14 subjects in 2016 and 12 regions in 2017 were again assigned to Cluster 1. The analysis of primary statistical information on the regions that made such a return allowed us to conclude that the main reason for it was a noticeable decrease in the level of real incomes of the population, observed from 2015 to 2017. The Samara Oblast is one of the most striking representatives of this group of regions: the population's real income decreased by 25% (from 14.1 to 10.5 thousand rubles at 2005 prices) in 2013–2017. At the same time, the birth rate of fourth and subsequent children showed a steady increase (from 0.046 to 0.062).

Cluster 2. In comparison with the previous cluster, in Cluster 2 the birth rate of fourth and subsequent children is lower, and the values of socio-economic indicators are higher. Average monetary income in this cluster in the first version

Table 1. Characteristics of clusters according to convergence of large families in Russian regions based on the analysis of 45 regions (2005–2017)

Variable	Average variable value by cluster				
	1	2	3	4	5
Birth rate of fourth and subsequent children	0.06488	0.06056	0.06331	0.06965	0.08533
Real monetary income by regional PPP, rubles per month	7 662	8 802	9 762	11 195	11 782
Nominal GRP per capita, rubles per year	152 369	162 941	228 583	336 077	517 927
Unemployment rate, %	7.6	6.9	6.2	5.7	5.7
Share of the population with income below subsistence minimum, %	20.4	15.6	14	12.6	13.2
Housing provision, sq. m per capita	22.9	23.2	24.2	24.8	24.9
R/P 10% ratio	11.5	12.8	13.3	13.9	14.2
Population per hospital bed, people	99.9	102.4	106.7	111.5	110.9
Coverage of children with preschool education, % of the number of children aged 1-6	60.8	62.8	65.9	69.2	72.1
Marriage rate	7.5	8.0	8.2	8.2	8.3
Life expectancy at birth, years	67.8	68.3	69	69.7	70.7
Migration growth, thousand people	-3.6	-6.0	-0.17	14	-5.0
Human Development Index	0.78	0.80	0.82	0.84	0.87
Percentage of urban population	66.2	70.2	72.1	74.7	77.7
Divorce rate	4.12	5.62	4.80	4.84	4.80
Number of abortions per 100 births, units	93.1	85.2	73.7	60.0	56.6
Calculated on the basis of the data from the Federal State Statistics Service of the Russian Federation.					

Table 2. Characteristics of clusters according to convergence of large families in 85 RF regions

Variable	Average variable value by cluster					
	1	2	3	4	5	6
Birth rate of fourth and subsequent children	0.095	0.084	0.071	0.079	0.101	0.194
Real monetary income by regional PPP, rubles per month	7 285	9 107	10 010	11 137	13 655	17 007
Nominal GRP per capita, rubles per year	94 752	182 605	268 030	418 157	1 280 278	5 115 483
Unemployment rate, %	10.3	7.3	6.0	5.7	4.5	5.6
Share of the population with income below subsistence minimum, %	21.4	15.9	13.7	13.2	10.9	7.8
Housing provision, sq. m per capita	21.2	23.4	25.0	25.2	23.9	24.2
R/P 10% ratio	11.8	12.9	13.2	13.6	14.7	16.6
Population per hospital bed, people	100.9	105.96	108.2	113.2	96.5	115.8
Coverage of children with preschool education, % of the number of children aged 1-6	55.7	62.6	67.6	70.7	74.7	78.0
Marriage rate	7.7	7.9	8.1	8.0	8.2	8.3
Life expectancy at birth, years	67.6	69.1	69.6	70.4	69.5	72.1
Migration growth, thousand people	-15.5	-10.0	-1.1	3.7	-22.2	-59.0
Human Development Index	0.77	0.81	0.84	0.86	0.89	0.89
Percentage of urban population	63.5	67.2	72.8	75.3	82.6	78.2
Divorce rate	4.0	4.5	4.8	4.7	5.3	4.6
Number of abortions per 100 births, units	94.1	71.2	64.0	56.5	67.1	56.5
Calculated on the basis of the data from the Federal State Statistics Service of the Russian Federation.						

of the calculations (see Table 1) amounts to 8,802 rubles, which is by 1,260 rubles (13%) higher than in Cluster 1, and rates of unemployment and poverty are lower – 6.93 and 15.6%, respectively. It is important to note a significantly lower average number of abortions (85.2 per 100 births). At the beginning of the period under analysis, the cluster included 9 regions with relatively high socio-economic indicators. In the period from 2006 to 2014, the cluster was replenished with subjects included in the sample for the first time, for example, regions of the Central Federal District (Vladimir, Lipetsk, Ryazan, Smolensk oblasts, etc.). In addition, regions from Cluster 1 were transferred to Cluster 2 due to the changes in socio-economic and demographic indicators. In 2012, Cluster 2 consisted of 17 regions that were in Cluster 1 at the beginning of the study period. The largest number of subjects in this cluster was in 2012–2013 (29), after which it gradually decreased due to the transition of regions to Cluster 3. In 2017, 10 regions remained in Cluster 2 (see Table 3).

Table 3. Distribution of RF subjects by cluster*

RF subject	Cluster membership												
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Central Federal District													
Belgorod Oblast	1	1	2	3	3	3	4	4	4	4	5	5	5
Bryansk Oblast	1	1	1	1	1	1	2	2	2	2	3	3	3
Vladimir Oblast	no data	no data	no data	no data	no data	no data	2	2	2	3	3	3	3
Voronezh Oblast	1	1	1	1	1	2	3	3	3	3	4	4	4

Continuation of Table 3

RF subject	Cluster membership												
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Ivanovo Oblast	1	1	1	1	1	1	1	1	2	2	2	2	2
Kaluga Oblast	1	1	1	2	2	2	3	3	3	4	4	4	4
Kostroma Oblast	1	1	1	1	1	2	2	3	3	3	3	3	3
Kursk Oblast	1	1	1	2	2	2	3	3	3	3	3	4	4
Lipetsk Oblast	1	2	2	3	2	3	3	3	3	4	4	4	4
Moscow Oblast	1	2	2	3	3	3	3	4	4	4	4	5	5
Orel Oblast	1	1	1	1	1	2	2	2	3	3	3	3	3
Ryazan Oblast	no data	no data	no data	no data	no data	no data	no data	2	3	3	3	3	4
Smolensk Oblast	no data	no data	no data	no data	no data	no data	no data	2	2	3	3	3	3
Tambov Oblast	no data	no data	no data	no data	no data	no data	no data	2	2	2	3	3	3
Tver Oblast	no data	no data	no data	no data	no data	no data	no data	no data	2	2	3	3	3
Tula Oblast	1	1	1	2	2	2	2	3	3	3	4	4	4
Yaroslavl Oblast	1	1	2	2	2	2	3	3	3	3	4	4	4
Moscow	no data	no data	no data	no data	no data	no data	no data	no data	5	5	5	5	5
Northwestern Federal District													
Republic of Karelia	**1**	**1**	**2**	2	2	2	3	3	3	3	4	4	4
Komi Republic	2	3	3	4	4	4	5	5	5	5	5	5	5
Arkhangelsk Oblast	1	2	3	3	3	3	4	4	4	5	5	5	5
Nenets AO	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	6	6
Vologda Oblast	no data	no data	no data	no data	no data	no data	no data	3	3	4	4	4	4
Kaliningrad Oblast	no data	no data	no data	no data	no data	2	3	3	3	4	4	4	4
Leningrad Oblast	1	2	2	3	3	3	4	4	4	4	5	5	5
Murmansk Oblast	2	2	3	3	3	3	4	4	4	4	5	5	5
Novgorod Oblast	1	1	1	2	2	3	3	3	3	4	4	4	5
Pskov Oblast	1	1	1	1	1	1	2	2	2	2	3	3	3
Saint-Petersburg	2	2	3	3	3	4	4	5	5	5	5	5	5
Central Federal District													
Republic of Adygea	1	1	1	1	1	1	1	2	2	2	3	3	3
Republic of Kalmykia	1	1	1	1	1	1	1	1	2	2	2	3	3
Republic of Crimea	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
Krasnodar Krai	no data	no data	no data	no data	no data	no data	no data	3	3	4	4	4	4
Astrakhan Oblast	1	1	no data	no data	no data	2	2	2	3	3	4	4	4
Volgograd Oblast	1	1	1	2	2	2	2	3	3	3	3	3	3
Rostov Oblast	no data	no data	no data	no data	no data	no data	2	2	2	3	3	3	4
Sevastopol	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data

Continuation of Table 3

RF subject	Cluster membership												
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
North Caucasian Federal District													
Republic of Dagestan	1	1	1	1	1	1	1	1	2	2	2	2	3
Republic of Ingushetia	1	no data	1	1	no data	no data	no data	1	1	1	1	1	1
Kabardino-Balkarian Republic	no data	no data	no data	1	1	1	1	1	1	2	2	2	2
Karachay-Cherkess Republic	no data	no data	no data	no data	1	1	1	1	2	2	2	2	2
Republic of North Ossetia	no data	no data	no data	no data	no data	no data	no data	2	2	2	2	2	2
Chechen Republic	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
Stavropol Krai	1	1	1	1	1	1	2	2	2	3	3	3	3
Volga Federal District													
Republic of Bashkortostan	1	2	2	2	2	3	3	3	3	4	4	4	1
Mari El Republic	1	1	1	1	1	2	2	3	3	3	3	1	1
Republic of Mordovia	1	1	1	1	2	2	2	3	3	3	3	1	2
Republic of Tatarstan	2	3	3	3	3	4	4	4	4	5	5	5	4
Udmurt Republic	1	2	2	2	3	3	3	3	4	4	4	1	1
Chuvash Republic	1	1	1	1	2	2	2	2	3	3	3	1	1
Perm Oblast	no data	no data	no data	no data	no data	no data	no data	4	4	4	4	4	4
Kirov Oblast	1	1	1	1	2	2	2	2	3	3	3	1	1
Nizhny Novgorod Oblast	no data	no data	2	2	no data	2	3	3	3	3	4	4	4
Orenburg Oblast	no data	no data	no data	no data	no data	2	3	3	4	4	4	4	4
Penza Oblast	1	1	1	1	2	2	3	3	3	3	3	1	2
Samara Oblast	2	3	2	3	3	3	4	4	4	4	4	1	1
Saratov Oblast	no data	no data	no data	no data	no data	no data	2	2	2	2	3	3	3
Ulyanovsk Oblast	1	1	1	2	2	2	3	3	3	3	3	1	1
Ural Federal District													
Kurgan Oblast	no data	no data	1	1	no data	no data	no data	2	2	2	2	2	3
Sverdlovsk Oblast	no data	no data	no data	no data	no data	3	3	4	4	4	4	4	4
Tyumen Oblast without AO	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
Khanty-Mansi Autonomous Okrug	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	5	5
Yamalo-Nenets Autonomous Okrug	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	6	6
Chelyabinsk Oblast	2	2	2	2	3	3	3	3	4	4	4	1	1
Siberian Federal District													
Altai Republic	no data	no data	no data	no data	no data	no data	no data	2	2	2	2	2	2
Tyva Republic	no data	no data	no data	no data	no data	no data	1	1	1	2	2	2	2
Republic of Khakassia	1	2	2	2	3	3	3	3	4	4	4	2	3

End of Table 3

RF subject	Cluster membership												
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Altai Krai	1	1	no data	no data	no data	no data	no data	2	2	2	2	2	2
Krasnoyarsk Oblast	3	3	3	4	4	4	5	5	5	5	5	1	1
Irkutsk Oblast	1	1	no data	no data	no data	no data	no data	3	4	4	4	4	4
Kemerovo Oblast	2	3	2	3	3	3	3	3	3	4	4	1	1
Novosibirsk Oblast	2	2	2	2	3	3	3	4	4	4	4	1	1
Omsk Oblast	2	2	2	2	3	3	3	3	4	4	4	2	2
Tomsk Oblast	3	3	3	3	4	4	4	4	5	5	5	2	3
Far Eastern Federal District													
Republic of Buryatia	no data	no data	no data	no data	no data	no data	no data	2	2	2	2	2	2
Republic of Sakha (Yakutia)	3	3	4	4	4	5	5	5	5	5	5	1	1
Zabaykalsky Krai	no data	no data	no data	no data	no data	no data	no data	2	2	2	2	2	2
Kamchatka Krai	no data	no data	no data	no data	no data	no data	no data	4	4	4	4	4	4
Primorsky Krai	1	2	2	3	3	3	3	4	4	4	4	1	2
Khabarovsk Krai	no data	no data	2	2	2	3	3	4	4	4	4	4	4
Amur Oblast	1	2	2	3	3	3	3	3	4	4	4	4	4
Magadan Oblast	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data	5	5
Sakhalin Oblast	no data	no data	no data	no data	no data	no data	5	5	5	5	5	5	5
Jewish Autonomous Oblast	no data	no data	no data	no data	2	2	2	3	2	3	3	3	4
Chukotka Autonomous Okrug	no data	no data	no data	no data	no data	no data	no data	5	5	5	5	5	5

* The distribution by cluster is presented in accordance with the available data on the birth rates of fourth and subsequent children in RF subjects.
Source: authors' calculations.

Cluster 3. It is characterized by the lowest average birth rate of fourth and subsequent children and the positive dynamics of socio-economic indicators in comparison with the previous clusters: real monetary income in the first variant is 9,762 rubles, in the second – 10,010 rubles, the unemployment rate accounts to 6.9 and 6.0%, respectively, the poverty rate – 14.0 and 13.7%, housing provision – 24.2 and 25.0 sq.m. Besides, the coverage of population with preschool education is going up. Cluster 3 was formed only in 2007 due to the changes in the socio-economic and demographic situation in RF subjects, as well as the inclusion of new regions in the sample. It comprised the Komi Republic, the Murmansk Oblast, Saint Petersburg and Krasnoyarsk Krai during the transition from Cluster 2. The Republic of Sakha (Yakutia) was included in the sample for the first time in 2007 and was also assigned to Cluster 3 based on the results of clustering. Then it was gradually expanded up to 2014 due to the transition of regions from Cluster 2. Its gradual, not-significant reduction has begun since 2015 (16 regions in 2017).

Cluster 4. This cluster is a turning point in some sense. The birth rate of fourth and subsequent children is gradually increasing and socio-economic indicators are improving: population's

real monetary income in the first variant amounts to 11,195 rubles, in the second – 11,137 rubles, the unemployment rate accounts for 5.7% in both variants, the poverty rate – 12.6 and 13.2%, respectively, and housing provision – 24.8 and 25.2 sq. m. In addition, the regions of this cluster alone show a positive migration increase and a minimum number of abortions per 100 births. The cluster formation began in 2010, and up to 2014 it expanded mainly due to changes in the socio-economic and demographic indicators of the studied regions (Komi Republic, Saint Petersburg, Krasnoyarsk Krai and Republic of Sakha (Yakutia)). Then, it gradually included regions from Cluster 3, as well as subjects, considered in the sample for the first time (Perm Krai and Kamchatka Krai). In 2014–2015, a steep growth of the cluster was recorded due to the transition of regions from Cluster 3. As a result, Cluster 4, comprising 28 regions, has become predominant by the end of the period under review.

Cluster 5. In Cluster 5 the trend that has appeared in Cluster 4 continues: there is a significant increase in the birth rate of fourth and subsequent children (in the second version of the calculations, it is 0.101, i.e. more than 25% higher than the Cluster 4 value), as well as a notable rise in socio-economic well-being: real monetary income in the first version is 11,782 rubles, in the second – 13,655 rubles, the unemployment rate is 5.7 and 4.5% (the lowest indicator among all clusters); and the poverty rate is 13.2 and 10.9%. This cluster has the largest share of urban population (82.6%). Cluster 5 emerged only in 2012 and expanded by including additional regions (Moscow, Khanty-Mansi Autonomous Okrug and Magadan Oblast) in the research database. In 2017, it comprised 13 regions.

Cluster 6. It is the smallest cluster (two regions only), sharply distinguished by the high birth rate of fourth and subsequent children (0.194), which is 92% higher than in Cluster 5. Cluster 6 also has high indicators of economic development: per capita monetary income amounts to 17,007 rubles, and the poverty rate is 7.8%. Nevertheless, there is also the highest migration outflow of the population – 59 thousand people. The cluster was formed only in 2016 due to the inclusion of additional regions (Nenets and Yamalo-Nenets autonomous okrugs) in the sample.

As a result, we can notice high dynamics of transitions of RF subjects between clusters in the period under review. Relatively high rates of economic development in Russia in 2005–2009 led to the transition of most regions from Cluster 1 to Cluster 2, characterized by a lower birth rate of fourth and subsequent children and higher indicators of socio-economic development. It should also be noted that the formation of Cluster 3 has begun since 2008, continuing the trend of moving from Cluster 1 to Cluster 2, with a decrease in the average birth rate of fourth and subsequent children to 0.071 and an increase in the values of socio-economic indicators. In 2008, Cluster 4 appeared (republics of Komi and Sakha (Yakutia)). We can say that it became a turning point in terms of demographics due to a rising birth rate of fourth and subsequent children and growing values of socio-economic indicators. From 2009 to 2017, most RF regions moved to Cluster 4.

In 2011, Cluster 5 (Sakhalin Oblast, since 2012 – Chukotka Autonomous Okrug) was formed. Possibly these regions could have previously been singled out into a separate cluster, but the earliest data on the birth rate of fourth and subsequent children appeared only in these years. There is a similar situation with Khanty-Mansi and Nenets autonomous okrugs, which joined Cluster 6 in 2016 (there are no data on these regions for earlier years).

It should be noted that out of 15 regions representing clusters 5 and 6 in 2017, only the Republic of Sakha (Yakutia) fell into this group as a result of development. According to the first observations results, it was assigned to Cluster 3, then, in 2008–2015 – to Cluster 4, and in 2016 –

to Cluster 5. The rest 14 regions were immediately placed in clusters 5 and 6, as soon as the necessary information on the birth of fourth and subsequent children appeared.

Conclusions

The cluster analysis allowed us to identify key trends in the regional differentiation of Russian large families based on birth rates of fourth and subsequent children and indicators of regions' socio-economic development. First, it is a decrease in the birth rate of many children with a one-time increase in the well-being of regions, which is confirmed by the dynamics of transition of RF subjects from Cluster 1 to clusters 2 and 3 in the studied period. Second, it is a trend associated with a rise in the birth rate of fourth and subsequent children with a simultaneous increase in well-being – transitions from Cluster 3 to clusters 4 and 5. It is necessary to note different intensity of this process. Many regions on this path made only one transition in 2012–2017: for example, the Ivanovo Oblast – from Cluster 1 to the subsequent one, the Tomsk Oblast – from Cluster 3 to Cluster 4. A number of subjects of the Russian Federation moved past two clusters. The transition from Cluster 1 to Cluster 3, characterized by grown welfare and a declined birth rate of fourth and subsequent children, was observed in 10 regions (Republics of Kalmykia and Mordovia, Kostroma, Orel, Pskov, Volgograd, Penza, Saratov and Kurgan oblasts, Stavropol Krai). In six subjects during the period under review we recorded growth of well-being, decline in the birth rate first and then its growth – these are transitions from Cluster 2 to Cluster 4 (republics of Komi and Tatarstan, Murmansk, Samara and Novosibirsk oblasts) and from Cluster 2 to Cluster 5 (Saint Petersburg). Only a small number of regions remained unchanged: Ingushetia remained in Cluster 1, while 3 Russian regions (republics of North Ossetia and Altai, and Altai Krai) remained in Cluster 2. Interesting were the results that showed the reverse movement of regions – from Cluster 1 to clusters 2 and 3, and then again to Cluster 1 (republics of Bashkortostan, Mari El, Udmurtia, and Chuvashia, Kirov and Ulyanovsk oblasts).

As a result, we get the following configuration of regional differentiation of large families and regions' socio-economic development: two extreme fairly stable "poles" – "the poor with many children" and "the rich with many children", between which there are other RF subjects, gradually shifting from the pole of "the poor with many children", but not yet approaching "the rich with many children". The regions with a high birth rate of fourth and subsequent children and low values of indicators of socio-economic development can be classified as "the poor with many children", while regions with high birth rates and high indicators of socio-economic development can be classified as "the rich with many children". As of 2017, the "pole" of "the poor with many children" includes 22 RF subjects from clusters 1 and 2: all the regions of the Caucasus, republics of Bashkortostan, Buryatia, Tyva, Chuvashia, and Altai, as well as a number of regions of the "middle zone" of Russia (Kirov, Ivanovo, and Ulyanovsk oblasts, etc.). The second group of "the rich with many children" consists of Russian regions belonging to clusters 5 and 6, including Moscow, Saint Petersburg, Moscow and Leningrad oblasts, as well as a number of regions of Siberia and the Far East (Khanty-Mansi and Yamalo-Nenets autonomous okrugs, Republic of Sakha (Yakutia), etc.).

The results obtained, in our opinion, refute several existing stereotypes about large families in Russia – its predominance among "poor" segments of the population and in rural areas. The study shows that the birth rate of fourth and subsequent children in 2017 was higher in the subjects of the Russian Federation with higher indicators of socio-economic development, where most population lives in cities. Accordingly, the efforts of state policy to increase the birth rate should be aimed

at moving from the model of large families with an "unplanned" nature of births, characteristic of low-income population groups [28], to the "conscious" and "planned" one. To do this, it is necessary to create a favorable life situation in which it becomes possible to implement "high" reproductive attitudes of Russian women. In this regard, increasing real incomes of the population, raising the level of socio-economic development of the regions located in clusters 3 and 4, which have capacity to move towards the pole of "the rich with many children", come to the fore.

References

1. Trynov A.V., Kostina S.N., Bannykh G.A. Examination of socio-economic determinants of fertility based on the regional panel data analysis. *Ekonomika regiona=Economy of Region*, 2020, vol. 16, no. 3, pp. 807–819. DOI: 10.17059/ekon.reg.2020-3-10 (in Russian).
2. Arkhangel'skii V.N. *Faktory rozhdaemosti* [Fertility Factors]. Moscow: TEIS, 2006. 399 p.
3. Bagirova A.P. *Rozhdaemost' i roditel'stvo v Rossii: determinanty i regional'naya differentsiatsiya: monografiya* [Fertility and Parenthood in Russia: Determinants and Regional Differentiation: Monograph]. Under general editorship of A.P. Bagirova. Yekaterinburg: Izd-vo Ural. un-ta, 2018. 157 p.
4. Istomina K.S. Impact of indicators on the birth rate in the regions. *Vestnik nauki i obrazovaniya= Bulletin of Science and Education*, 2015, vol. 4, no. 2 (4), pp. 60–63 (in Russian).
5. Koreva O.V., Boitsova T.E. Analysis and evaluation of the impact of individual factors on the state of the modern demographic situation in the Russian Federation. *Vestnik evraziiskoi nauki=The Eurasian Scientific Journal*, 2013, vol. 19, no. 6, pp. 47–59. (In Russian).
6. Sukneva S.A., Barashkova A.S., Postnikova K.Yu. Birth rate, number of children and family income: trends and relationships. *Ekonomicheskie i sotsial'nye peremeny: fakty, tendentsii, prognoz=Economic and Social Changes: Facts, Trends, Forecast*, 2020, vol. 13, no. 2, pp. 201–213. DOI: 10.15838/esc.2020.2.68.13 (in Russian).
7. Zhuravleva T.L., Gavrilova Ya.A. Analysis of fertility determinants in Russia: What do RLMS data say? *Ekonomicheskii zhurnal VShE=HSE Economic Journal*, 2017, vol. 21, no. 1, pp. 145–187 (in Russian).
8. Mietule I., Maksymova I., Holikova K. Modern socio-demographic trends affecting the business sector of the economies of Latvia and Ukraine. In: *Society. Integration. Education. Proceedings of the International Scientific Conference*. 2018. Vol. 6. Pp. 353–365.
9. Boyle P.J., Graham E., Feng Z. *Contextualising Demography: The Significance of Local Clusters of Fertility in Scotland*. Max Planck Institute for Demographic Research, Rostock, Germany, 2007. DOI: 10.4054/MPIDR-WP-2007-036
10. Salvini M.S., Gabrielli G., Paterno A., Corazziari I. Demographic trends in developing countries: convergence or divergence processes? In: *Corazziar Disia Working Paper*. 2015. No. 03. Available at: https://local.disia.unifi.it/wp_disia/2015/wp_disia_2 (accessed: November 10, 2020).
11. Jurun E., Ujević I. A cluster analysis of Croatian counties as the base for an active demographic policy. *Croatian Operational Research Review*, 2017, vol. 8, no. 1, pp. 221–236. DOI: https://doi.org/10.17535/crorr.2017.0014. Available at: https://hrcak.srce.hr/181659
12. Pénzes J., Pásztor I.Z., Tátrai P. Demographic processes in developmentally peripheral areas of Hungary. *Stanovnistvo*, 2015, vol. 53, no. 2, pp. 87–111. DOI: 10.2298/STNV1502087P
13. Yüceşahin M., Tulga A. Demographic and social change in the middle East and North Africa: Processes, spatial patterns, and outcomes. *Population Horizons*, 2017, vol. 14, no. 2, pp. 47–60. DOI:10.1515/pophzn-2017-0003/
14. Petrykina I.N. Cluster analysis of regions of the Central Federal District in terms of human capital development. *Vestnik VGU. Seriya: Ekonomika i upravlenie=Proceedings of Voronezh State University. Series: Economics and Management*, 2013, no. 1, pp. 72–80 (in Russian).
15. Filipova A.G., Es'kova A.V., Inzartsev A.V. Social potential of a region: Experience of using cluster analysis. *Regionologiya=Russian Journal of Regional Studies*, 2017, vol. 100, no. 3, pp. 438–455 (in Russian).

16. Kanishchev V.V Cluster analysis of demographic behavior of the rural population of European Russia at the beginning of the 20th century and at the beginning of the 21st century. On raising the question. *Ineternum= Ineternum*, 2011, no. 1, pp. 43–55 (in Russian).
17. Shubat O.M., Bagirova A.P., Akishev A.A. Methodology for analyzing the demographic potential of Russian regions using fuzzy clustering. *Ekonomika regiona=Economy of Region*, 2019, vol. 15, no. 1, pp. 178–190. DOI: 10.17059/2019-1-14 (in Russian).
18. Golubova T.N., Makhkamova Z.R., Ovsyannikova N.M. Cluster analysis of birth and mortality rates in the Republic of Crimea. *Nauchnye vedomosti. Seriya: Meditsina. Farmatsiya=Belgorod State University Scientific bulletin. Medicine. Pharmacy*, 2016, vol. 34, no. 12, pp. 88–94 (in Russian).
19. Russkikh T.N., Tinyakova V.I., Stroev S.P. Typologization of Russian regions by efficiency level of health care systems using fuzzy clustering procedures. *Uchet i statistika= Accounting and Statistics*, 2017, vol. 47, no. 3, pp. 43–52 (in Russian).
20. Dolgodvorova E.V. Cluster analysis: Basic concepts and algorithms. *Voprosy nauki i obrazovaniya=Questions of Science and Education*, 2018, vol. 19, no. 9, pp. 73–76 (in Russian).
21. Kovanova E.S. Cluster analysis in handling the problem of typology of regions of Russia by level and intensity of internal labor migration. *Vestnik NGUEU= Vestnik NSUEM,* 2013, no. 4, pp. 166–175 (in Russian).
22. Wijayanto A.W., Purwarianti A., Son L.H. Fuzzy geographically weighted clustering using artificial bee colony: An efficient geo-demographic analysis algorithm and applications to the analysis of crime behavior in population. *Appl Intell*, 2016, vol. 44, pp. 377–398. DOI:/10.1007/s10489-015-0705-7
23. Le Hoang Son. A novel kernel fuzzy clustering algorithm for geo-demographic analysis. *Inf. Sci*, 2015, vol. 317, pp. 202–223. DOI: 10.1016/j.ins.2015.04.050
24. Shubat O.M., Shmarova I.V. Cluster Analysis as an Analytical Tool of Population Policy. *Ekonomika regiona=Economy of Region*, 2017, vol. 13, no. 4, pp. 1175–1183 (in Russian).
25. Ilyshev A.M., Bagirova A.P. Factors of reproductive activity of the population: Analysis of regional differentiation. *Region: ekonomika i sotsiologiya=Region: Economics and Sociology,* 2009, no. 1, pp. 92–110 (in Russian).
26. Tikhomirov N.P., Tikhomirova T.M. Methods of justification for strategies to overcoming demographic crisis in the regions of Russia. *Fundamental'nye issledovaniya=Fundamental research*, 2020, no. 5, pp. 160–166 (in Russian).
27. Ward J.H. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 1963, pp. 236–244.
28. Kostina S.N., Zaitseva E.V. Large families models in modern Russian society: The results of narrative interviews. *Sotsiologicheskie issledovaniya=Sociological Studies,* 2021, no. 3, pp. 92–102. DOI: 10.31857/S013216250009563-2 (in Russian).

Information about the Authors

Svetlana N. Kostina – Candidate of Sciences (Sociology), Associate Professor, Associate Professor of the Department, Ural Federal University named after the First President of Russia B.N. Yeltsin (19, Mira Street, Yekaterinburg, 620002, Russian Federation; e-mail: s.n.kostina@urfu.ru)

Aleksandr V. Trynov – Junior Researcher, Institute of Economics, Ural Branch of the Russian Academy of Sciences (29, Moskovskaya Street, Yekaterinburg, 620014, Russian Federation; e-mail: Trynovv@mail.ru)

Received November 17, 2020.