The 14th International Days of Statistics and Economics, Prague, September 10-12, 2020

THE STUDY OF SPATIAL AUTOCORRELATION OF FERTILITY IN RUSSIA

Oksana Shubat – Grigory Ivanenko – Mark Shubat

Abstract

Russia has historically developed a high level of regional differentiation, including in terms of population reproduction. The aim of the study is to assess the spatial autocorrelation of fertility in Russia. Our analysis uses the data on regional fertility levels for 2000-2015, which was the last period of sustained fertility growth in Russia. To explain spatial effects, we applied Moran's I calculated by using three different spatial weights matrices and Barro's regression (the concept of β-convergence). Our results did not show that neighbouring Russian regions form clusters based on fertility levels and fertility trends. Moran's indices calculated by using three different spatial weights matrices produced contradictory results regarding the level of geographical clustering of Russian regions. We also have not found any evidence that there are convergence trends in fertility rates of Russian regions unrelated to their geographical proximity. Parameters of Barro's regression have proven to be statistically significant but the explanatory power of the equation is extremely low. Thus, our findings do not show that the diffusion theory of fertility is applicable for country-specific demographic analysis in Russia.

Key words: fertility, Moran’s index, Barro regression, Russian regions

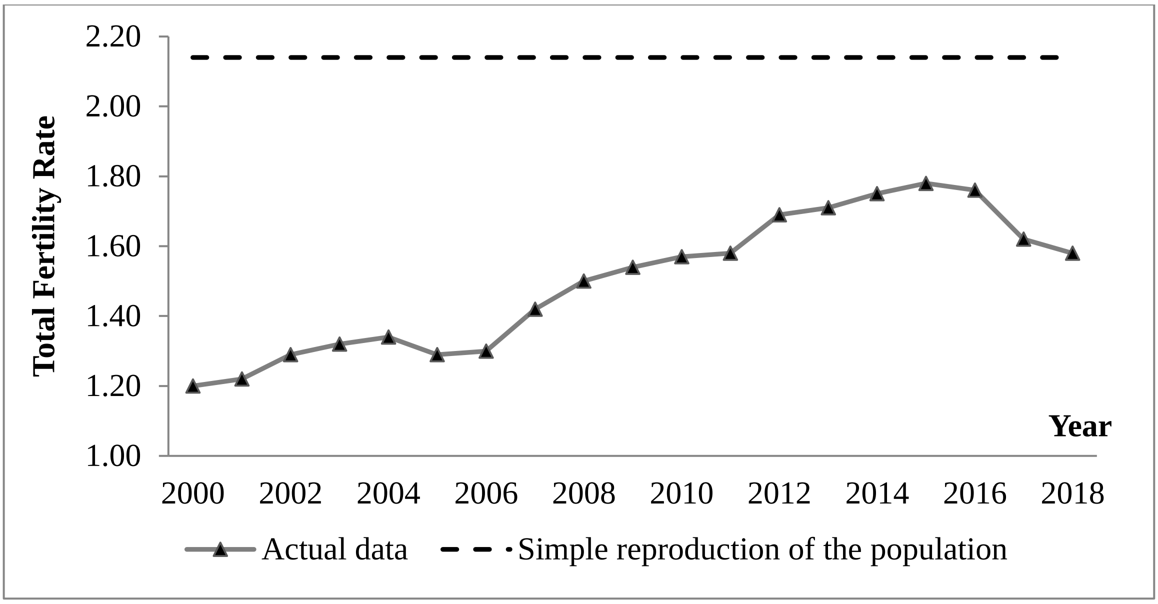
JEL Code: C12, J13

Introduction

Russia has historically developed a high level of regional differentiation, including in terms of population reproduction. For instance, in 2018, the crude birth rate in Russian regions varied from 7.6 to 20.7; the crude death rate varied from 3.1 to 17.2; and the total fertility rate, from 1.124 to 2.967 (Total Fertility, 2020). It is important to take into account this differentiation to evaluate the prospects of demographic development and devise efficient policy solutions in this sphere.

The question of state supported reproduction, such as subsidies and financial incentives, is important for Russia. The country has already reached belowreplacement fertility level and the fall in the fertility rates has continued in the recent years (Fig.1), despite the government's efforts to reverse this trend. Spatial analysis of fertility patterns is necessary to devise effective measures for boosting fertility on the state level.

Fig. 1: Total fertility rate in Russia



Source: compiled by the author on the basis of (Total Fertility, 2020)

Spatial autocorrelation of fertility in Russia remains a largely underexplored question. Research on this topic is relatively scarce not only for Russia but for other countries as well. The main theory that underlies such research is the diffusion theory of fertility (see, for example, (Cleland & Wilson, 1987)), which explains the massive decline in fertility during the demographic transition by the changing ideas about the ideal family size, the appearance of new methods of fertility control, and so on.

Due to the peculiarities of territorial differences in fertility levels within some countries, it is also possible to apply the diffusion theory of fertility for country-specific analysis. In this case the changing fertility levels are explained the following way: the transformations affect the most advanced regions first and then spread to peripheral areas. For example, Vitali and Billari (Vitali & Billari, 2015) showed that the diffusionist perspective to fertility transition could still be relevant in explaining fertility changes in Italy. Carioli et al. (2014) found a strong spatial autocorrelation of fertility in Spain. Inozemcev and Kochetygova (Inozemcev & Kochetygova, 2018) confirmed the existence of spatial effects for fertility in some Russian regions. In some studies, the diffusion of fertility is considered even at the level of individuals and families. For instance, Balbo and Barban (Balbo &

Barban, 2014) demonstrated that a friend's childbearing increases an individual's risk of becoming a parent.

Another aspect of spatial analysis of fertility patterns focuses on convergence in regional fertility rates. In research literature there are several approaches to convergence analysis. Barro and Sala-I-Martin describe the concepts of σ-convergence, absolute and conditional β-convergence (Barro & Sala-I-Martin, 1992). There are, however, relatively few studies of regional fertility convergence. For example, Franklin and Plane analyze β- and σ-convergence as well as club convergence of regional fertility rates in Italy (Franklin & Plane, 2003). In our previous study, we also examined the convergence of regional fertility in Russia (Shubat, 2018).

The purpose of the current study is to analyze spatial autocorrelation of fertility in Russian regions. To this end, we tested three hypotheses: 1) in Russia, geographically close regions form clusters based on their fertility rates; 2) geographically close regions form clusters based on their fertility growth rates; and 3) the regional fertility levels demonstrate convergent trends unrelated to the geographical proximity of regions.

1 Data and Methods

Our analysis relies on the data on fertility rates in Russian regions for 2000-2015, which was the last period of sustained fertility growth in the country. Since 2016, fertility rates in Russia have been declining. We used the total fertility rate (TFR) as an indicator of the fertility level.

The first hypothesis was tested by using the regional TFR data. The second hypothesis was tested by using the TFR growth rates in Russian regions. We analyzed two coefficients: the first showed fertility rates until the end of 2007 and the second corresponded to the period after 2007. The year of 2007 was chosen as a turning point since it was at this time that the Russian government launched an active campaign to stimulate fertility. This fact led us to suppose that spatial effects may be different in these two periods: from 2000 to 2007 and from 2007 to 2015.

To explain spatial effects (and to test the first and second hypotheses), we used Moran's I, which is the most widely used measure of spatial autocorrelation (Moran, 1950). The significance of the index is evaluated with the help of z-score while the type of spatial autocorrelation was determined by applying a standard approach based on comparison of the observed index values with the expected ones (Introduction to, 2006):

- if Im > E(I), the spatial autocorrelation is positive, that is, neighbouring regions have similar values of fertility rates (for the first hypothesis) and similar fertility growth rates (for the second hypothesis);
- if Im < E(I), then the spatial autocorrelation is negative, that is, the fertility levels (for the first hypothesis) and fertility growth rates (for the second hypothesis) in neighbouring regions are different;
- If Im = E(I), the fertility levels (for the first hypothesis) and fertility growth rates (for the second hypothesis) in neighbouring regions are distributed in a random manner.

The value of Moran's index is known to depend on the assumptions built into the spatial weights matrix. Construction of appropriate weight matrices has long proven to be a controversial topic in spatial modelling. To obtain more valid results, we calculated Moran's I by using three different spatial weights matrices.

In the first case, we used a binary matrix:

$$w_{ij} = \begin{cases} 1 & \text{if regions } i \text{ and } j \text{ share a common border);} \\ 0 & \text{if } i=j; \\ 0 & \text{if region } i \text{ does not share a common border with region } j. \end{cases}$$

In the second case, the matrix was formed on the basis of row standardization:

$$w_{ij(s)} = w_{ij} / \sum w_{ij} \tag{1}$$

In the third case, we applied another approach and formed a distance-based spatial weights matrix (inverse distance matrix). In order to determine a spatial weight, we used the shortest road distance between regional centers.

To test our third hypothesis about convergence in fertility rates in Russian regions unrelated to their geographical proximity, we estimated Barro's regression (Barro and Sala-I-Martin, 1992):

Barro regression is estimated as follows:

$$\frac{1}{T}\ln\left(\frac{y_{i,t+T}}{y_{i,t}}\right) = \alpha + \beta \cdot \ln y_{i,t} + \varepsilon_{i,t}, \tag{2}$$

where $y_{i,t}$ and $y_{i,t+T}$ are the fertility rates in region i in 1999 and 2015 respectively; $\frac{1}{T}\ln\left(\frac{y_{i,t+T}}{y_{i,t}}\right)$ is the average annual growth rate of fertility in region *i* in 2010-2015.

If the regression coefficient is statistically significant and is less than zero, then the hypothesis of β-convergence is confirmed.

2 Results

1. In the period from 2000 to 2015, there was an increase in fertility. The fertility rates, however, varied considerably across the Russian regions (Table 1). Variability in the regional fertility levels changed with time but in a non-linear way: in 2002-2002, it fell, then until 2008 it increased and then started to go down again.

Tab. 1: Total fertility rate in Russia

Year	TFR	Minimum TFR in regions	Maximum TFR in regions	Year	TFR	Minimum TFR in regions	Maximum TFR in regions
2000	1.20	0.93	2.46	2008	1.50	1.10	3.44
2001	1.22	0.98	2.17	2009	1.54	1.16	3.41
2002	1.29	1.03	2.10	2010	1.57	1.17	3.45
2003	1.32	1.08	2.29	2011	1.58	1.16	3.36
2004	1.34	1.10	2.99	2012	1.69	1.22	3.35
2005	1.29	1.02	2.95	2013	1.71	1.23	3.42
2006	1.30	1.01	2.81	2014	1.75	1.28	3.49
2007	1.42	1.06	3.18	2015	1.78	1.29	3.39

Source: Fertility data of the Russian single inter-departmental information and statistical system (Total Fertility, 2020)

2. Moran's I calculated by using the fertility data for all the given years exceeded the expected index value E(I) = - 0.012 (Table 2), which shows a positive spatial correlation. Thus, neighbouring regions tend to have similar fertility rates. For all the years except for 2000 and 2001, the indices proved to be statistically significant (p-value < 0.05).

Our analysis has shown that values of the indices change considerably if we change the type of the spatial weights matrix. The inverse distance matrix gives the highest values: in different years the values of Moran’s I we obtained by using the matrix of this type exceeded the indices calculated by using the binary matrix by 9-159% (Table 2). Importantly, the indices characterize changes in spatial clustering differently. The dynamics of the indices based on the row-standardized matrix and the distance matrix is very similar and different from the dynamics of the indices calculated on the binary matrix.

These outcomes are not conclusive, therefore, our first hypothesis that neighbouring regions in Russia form clusters with similar fertility levels cannot be confirmed with certainty.

Tab.2: Spatial autocorrelation values of regional fertility levels in Russia

Year	Moran's I calculated by using:			Ratio of Moran's I calculated by using an inverse distance matrix to:	
	binary matrix	row-standardized matrix	inverse distance matrix	Moran's I calculated by using a binary matrix	Moran's I calculated by using a row-standardized matrix
2000	0.119	0.310	0.309	2.595	0.996
2001	0.142	0.339	0.346	2.441	1.020
2002	0.182	0.389	0.403	2.210	1.034
2003	0.197	0.407	0.427	2.169	1.049
2004	0.264	0.490	0.591	2.236	1.206
2005	0.269	0.494	0.593	2.204	1.202
2006	0.286	0.530	0.635	2.220	1.198
2007	0.300	0.516	0.606	2.024	1.176
2008	0.309	0.519	0.611	1.979	1.177
2009	0.308	0.521	0.612	1.987	1.175
2010	0.293	0.498	0.589	2.013	1.184
2011	0.303	0.501	0.584	1.926	1.166
2012	0.329	0.557	0.637	1.934	1.144
2013	0.346	0.579	0.658	1.901	1.136
2014	0.396	0.423	0.444	1.119	1.048
2015	0.362	0.381	0.395	1.092	1.037
Increase in fertility rates from 1999 to 2007	0.204	0.507	0.630	3.096	1.243
Increase in fertility rates from 2007 to 2015	0.258	0.667	0.825	3.196	1.237

Source: author’s calculations.

3. The values of Moran's I calculated by using the data on the increase in fertility rates before and after 2007 exceeded the expected value of index E(I) = - 0.012, which means that neighbouring regions are more prone to demonstrating similar dynamics of fertility rates. However the indices calculated by using different spatial weights matrix provide opposite estimates of the geographical clustering of fertility rates; there are several-fold differences between the indices – from 1.2 to 3.2 times (Table 2). Thus, the second hypothesis that neighbouring regions form clusters with similar fertility trends cannot be confirmed.

4. To test our third hypothesis about convergent trends in regional fertility levels unrelated to the geographical proximity of regions, we estimated Barro's regression model

(*β*-convergence). Prior to that, we excluded two regions as extreme outliers in terms of statistical modelling (more than 3 interquartile range distance from the third quartile). Therefore, these regions should be excluded from the set of regions. The main parameters of the regression equations are presented in Tables 3-4.

Tab. 3: Model summary

R Square	Adjusted R Square	Std. Error of the Estimate	F	Sig.
0.259	0.250	0.005	27.622	0.000

Source: author's calculations.

Tab. 4: Coefficients

Model		Unstandardized Coefficients		t	Sig.
		B	Std. Error		
1	Constant	0.030	0.001	31.286	0.000
	Ln TFR_{1999}	-0.022	0.004	-5.256	0.000

Source: author's calculations.

As these data show, there is no convergence observed in the regional fertility levels. On the one hand, the parameters of the equations are statistically significant and parameter *β* is negative. Thus, regions with initially lower fertility levels seem to be "catching up" with regions with an initially higher level due to higher annual growth rates. At the same time, the explanatory power of the equations is rather low (does not exceed 26%). Therefore, the third hypothesis cannot be confirmed with certainty.

3 Discussion

Our research findings raise a number of questions. The first question to be considered in this respect is the contradictory results we obtained when calculating Moran's indices with the help of different spatial weights matrices. It should be noted at this point that the question about the most suitable type of matrix for studying spatial effects of fertility still remains open. There are no convincing demographic theories that would provide a clue as to what set of spatial weights produces the spatial autocorrelation of fertility, which determines internal relationships between spatial objects in terms of fertility. Anselin and Florax (1995) showed that calculation of Moran's I by using the set of weights different from the one that actually produces the spatial correlation leads to incorrect testing results.

More research cases using different spatial weights matrices are probably needed in order to gather enough empirical evidence to make a proper choice of the spatial weights matrix. In Russian research practice, though scarce, different matrices were used to study spatial effects of fertility, which brought different outcomes. Grigoriev obtained Moran's I value 0.48 for regional crude birth rates in 2012 (Grigoriev, 2018). Inozemcev and Kochetygova analyzed the TFRs in 55 Russian regions between 2004 and 2015 and calculated Moran's I by using three different spatial weights matrices. Their resulting values varied from 0.1 to 0.4. In their study, the indices calculated by using different matrices were different for each given year (Inozemcev & Kochetygova, 2018).

We cannot rule out the possibility that the fertility diffusion can be country-specific or particular. In this case, the use of the same spatial weights matrix can be effective for one country (or group of countries) but fail to bring any meaningful results for other countries.

The lack of convincing empirical evidence of convergent trends in regional fertility levels (as our testing of the third hypothesis based on the concept of β-convergence has shown) raises new questions. First of all, these results do not support the demographic transition theory, which the idea of fertility convergence is a part of. The applicability of this theory to the study of demographic processes in countries with considerable regional differences in terms of fertility rates is a debatable question.

Conclusions

The results of our analysis lead us to the following conclusions regarding the spatial autocorrelation of fertility rates in Russian regions. First, the hypothesis that in Russia neighbouring regions form clusters depending on their fertility rates cannot be confirmed with certainty. The values of Moran's I calculated by using different spatial weights matrices produced contradictory results regarding the degree of geographical clustering of Russian regions. Second, the hypothesis that Russian neighbouring regions form clusters depending on the fertility growth rates cannot be confirmed either. For these indicators, Moran's indices calculated by using three different spatial weights matrices brought contradictory results. Third, the hypothesis about convergent trends in regional fertility levels unrelated to the geographical proximity of regions cannot be confirmed with confidence. There is still a considerable imbalance between Russian regions in terms of fertility levels.

More demographic and statistical research is necessary to accumulate enough empirical data on spatial effects in regional fertility levels and to show that the diffusion theory of fertility transition may be adapted for studying country-specific (or regional) peculiarities of population reproduction. Our findings do not show that the diffusion theory is applicable for country-specific demographic analysis in Russia.

Acknowledgment

The article is one of the outputs of the research project “Russian pro-natalist policy: resources, effects, optimization opportunities“, supported by the Council on grants of the President of the Russian Federation, project no. NSh-2722.2020.6

References

Anselin L., & Florax R.J.G.M. (1995) New Directions in Spatial Econometrics: Introduction. In: Anselin L., Florax R.J.G.M. (eds) *New Directions in Spatial Econometrics. Advances in Spatial Science*. Springer, Berlin, Heidelberg

Balbo, N., & Barban, N. (2014). Does Fertility Behavior Spread among Friends? *American Sociological Review*, *79*(3), 412–431. doi: 10.1177/0003122414531596

Barro, R., & Sala-I-Martin, X. (1992). Convergence. *Journal of Political Economy*, *100*(2), 223-251. doi:10.1086/261816

Carioli, A., Devolder, D., & Recano, J. (2014). A spatial analysis of recent fertility patterns in Spain. *European Population Conference*, 25–28 June 2014. Budapest, Hungary. Retrieved from https://epc2014.princeton.edu/papers/140253

Cleland, J., & Wilson C. (1987). Demand Theories of the Fertility Transition: An Iconoclastic View. *Population Studies*, *41*(1), 5-30;

Franklin, R. S., & Plane, D. (2004). A Shift-Share Method for the Analysis of Regional Fertility Change: An Application to the Decline in Childbearing in Italy, 1952-1991. *Geographical Analysis*, *36*(1), 1-20. doi:10.1353/geo.2003.0021

Grigoriev, A. A. (2018). Spatial Autocorrelation of Educational Attainment in the Russian Federation. *Psychology. Journal of Higher School of Economics, 15*(1), 164–173. doi: 10.17323/1813-8918-2018-1-164-173.

Inozemcev, E. S., & Kochetygova O. V. Spatial Panel Analysis of Fertility and Life Expectancy in Russia (2018). *Izv. Saratov Univ. (N. S.), Ser. Economics. Management. Law*, 18(3), 314–321.

Introduction to Spatial Analysis. Invited Lecture. Population Science and GIS Workshop, UC Santa. 2006.

Moran, P. A. P. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika*. *37* (1): 17–23.

Shubat, O. (2018) Fertility convergence at the regional level: empirical evidence from Russia. *The 12th International Days of Statistics and Economics*. Retrieved from https://msed.vse.cz/msed_2018/article/32-Shubat-Oksana-paper.pdf

Total Fertility Rate data. Russian single inter-departmental information and statistical system (2020). Retrieved March 25, 2020, from https://fedstat.ru/indicator/31517

Vitali, A., & Billari, F. C. (2015). Changing Determinants of Low Fertility and Diffusion: a Spatial Analysis for Italy. *Population, Space and Place, 23*(2). doi: 10.1002/psp.1998

Contact

Oksana Shubat
Ural Federal University
Mira street, 19, Ekaterinburg, Russia, 620002
o.m.shubat@urfu.ru

Grigory Ivanenko
Ural Federal University
Mira street, 19, Ekaterinburg, Russia, 620002
ivanenkogrig@yandex.ru

Mark Shubat
Ural Federal University
Mira street, 19, Ekaterinburg, Russia, 620002
Mark.Shubat@urfu.me