Structural Change and Economic Dynamics 55 (2020) 244–258

Contents lists available at ScienceDirect

Structural Change and Economic Dynamics

journal homepage: www.elsevier.com/locate/strueco

Marshallian vs Jacobs effects: Which is stronger? Evidence for Russia unemployment dynamics

Olga Demidova[a,*], Alena Kolyagina[a], Francesco Pastore[b]

[a] National Research University Higher School of Economics, Moscow, Russia
[b] University of Campania "Luigi Vanvitelli", Italy

ARTICLE INFO

Article history:
Received 10 January 2020
Revised 15 July 2020
Accepted 30 July 2020
Available online 10 September 2020

JEL codes:
C14
C21
J64

Keywords:
Concentration
Diversification
Unemployment
Spatial effects
Nonparametric models

ABSTRACT

This paper studies the influence of diversification and specialization on one of the main indicators of the Russian labour market: unemployment growth. The purpose of the work is to find out which effects dominate in the Russian regions, Marshallian or Jacobs, and whether this predominance is stable for different time periods. We tested empirically the following hypotheses: 1) the dependence of the unemployment growth on the concentration or diversification is nonlinear due to possible overlapping effects of urbanization and localization; 2) the influence of the concentration or diversification on the unemployment growth depends on the time period. To test these hypotheses, we use nonparametric additive models with spatial effects. Both hypotheses found empirical confirmation, with each effect prevailing in different time periods: Marshallian effects were prevalent in 2008-2010, and 2013-2016, while Jacobs effects were prevalent in 2010-2013.

1. Introduction

Better knowledge of the differences between Russian regions allows the state to pursue a more structured national and regional policy in order to avoid the negative social and economic consequences of the concentration of regions with high unemployment (Elhorst, 2003). One of the most important factors of inequality is the current concentration of economic activities in regions which have a number of competitive advantages. The possible consequences of this high concentration and its impact on unemployment growth are important because of the existence of two effects with opposite signs. Jacobs theory (Jacobs, 1969) posits that due to higher diversification, urban territories absorb unemployment shocks better: in fact, it is easier to find a job in another sector of the economy in case of job loss, which leads to a lower unemployment rate. Marshall's theory, by contrast, suggests that regions with a high level of specialization have better economic indicators and have lower unemployment due to agglomeration economies (Marshall, 1920). In other words, local agglomeration of firms in one industry creates a labour market with a limited set of skills required for this particular industry. Labour resources contribute to the growth of productivity and the reduction of differences in wages during the transition from one employer to another. These effects can overlap, especially in heterogeneous regions, and the main objective of the study is to empirically confirm these effects; to find out, which effect dominates; and whether this predominance is constant during different time periods. Additionally, we test the applicability of models with a non-parametric component that works well for European data to model labour market indicators (particularly, unemployment growth) for Russian regions, and justify their advantage over simple linear models.

One should understand the agglomeration effect as the economic benefit deriving from the concentration of firms in a certain territory. Within the borders of agglomeration, it becomes possible to save costs for the interacting companies due to close cooperation if certain regions attract manufacturing factors (technologies, labour resources and investments). The agglomeration effect contributes to the emergence of competitive clusters, which, in turn, is an incentive for their concentration in a certain territory (Rastvortseva and Kuga, 2012).

Alfred Marshall was the first to notice the existing inclination of industries to territorial agglomeration, which contributes to the growth of profitability and economies of scale. According

* Corresponding author at: 11 Pokrovsky Bulvar, Pokrovka Complex, room S534, Moscow 109028, Russia.

E-mail addresses: demidova@hse.ru (O. Demidova), 4eely94@mail.ru (A. Kolyagina), francesco.pastore@unicampania.it (Pastore F.).

https://doi.org/10.1016/j.strueco.2020.07.010
0954-349X/

to Marshall, workers periodically change their place of work (among those that use this particular kind of labour), which makes it possible to increase productivity and reduce differences in wages. As a result of this mobility, workers are able to borrow knowledge and skills from each other in industrial clusters in a short period of time, and enterprises have the opportunity to recruit trained employees with ready to use knowledge and skills, which reduces the cost of training staff inside the company. Duranton and Puga (2014) provide more details about theoretical foundations of urban agglomeration, Combes, Gobilon (2015) consider the empirics. Parr (2012) discusses the coexistence of different types of agglomeration economies.

In contrast, Jane Jacobs believed that as the diversification of industries increases, the number of job opportunities for the population increases too, which leads to a reduction in the regional unemployment rate (Jacobs, 1969). The various interrelationships between large diversified cities allow the creation and implementation of innovations, which contribute to increased productivity and economic growth of each of the enterprises in a given territory. These effects were named Jacobs effects.

The disproportions in the spatial development of regions can be explained with the help of these theories of spatial distribution. There may be agglomeration effects from localization (under the Marshallian externalities), contributing to a reduction in production costs due to economies of scale, but the existence of centrifugal processes is also possible due to excessive infrastructure congestion, environmental problems, high population density, and increased transportation costs. The total agglomeration effect, which determines the concentration of production in the industry in any limited territory, is of particular interest.

The case of Russia is important for several reasons. First, the vast territories of the country provide evidence of very different and varied experiences of both agglomeration and diversification. This makes Russia a unique testing ground for Jacob versus Marshallian effects. In addition, the historical stratification of industry localizations makes several regions of the country traditionally specialized in specific types of industries as a consequence of the past forced industrialization. Partly, agglomeration economies are also linked, at least initially, to the localization of natural resources, especially gas and oil, and the relative mining industry. On the other hand, over the last three decades, the “disorganization” of central planning (Blanchard and Kremer, 1997) changed the past specialization pattern of several regions of the country, breaking down old linkages between industries and, therefore, generating a higher degree of diversification of productions especially in urban areas and new product specializations overlapping with the old ones in other less urbanized areas (for an analysis of the impact of industry diversification on the quality evolution of jobs, see (Gimpelson and Kapeliushnikov, 2016). Understanding the impact of the two effects on unemployment growth is important for policy makers interested in shaping future decisions regarding investment localization and in forecasting the impact of possible economic crises in specific sectors on employment outcomes, considering also the fragility over time of the Russian labor market, with high wage flexibility and rigid employment rates (Gimpelson and Kapeliushnikov, 2016; Voskoboynikov, 2017). Is this model bound to persist? What is the role of agglomeration factors in shaping it? If this model becomes infeasible what would be the employment consequence of this change with the occurrence of structural change? This paper aims to address directly or indirectly these types of questions.

This study innovates on previous research in several respects. First, we use Russian regional data over a relatively long period of time (2007-2016), which allows us to emphasize the possibility to test for differences in the effects from one period to the following period. In particular, thanks to our data, we are able to test whether there is a different dependence relationship in "crisis periods" and in more favorable periods. In fact, we find different effects in periods of ups and downs. Second, we used the Vorobiev index, which better describes the degree of regional diversification compared with the Herfindahl-Hirschman indices.

Third, to test the robustness of our findings, we also used Ellison-Glaser index of concentration. Additionally, indices of regional diversification and concentration of production were calculated in two ways using firm level data: on the basis of the revenues of companies and on the basis of value added by economic activity.

Fourth, we took into account the mutual influence of Russian regions by including spatial lags in our models. Fifth, we used flexible semiparametric dependence for each variable and ANOVA test for the choice between linear and nonlinear functional form.

The structure of the paper is as follows. In the next section, we provided a brief review of the papers highlighting the impact of the Jacobs and Marshallian effects in different countries. Section three presents our data source, the choice of the explanatory variables and the hypotheses to be tested. Section four describes the methodology of econometric modeling and presents the results of the estimation and their interpretation. The last section contains some concluding remarks and policy implications.

2. Literature review

There are many empirical works devoted to the influence of the concentration or diversification of a region on changes in unemployment in it. Simon (1988), Elhorst (2003), Ferragina and Pastore (2008) empirically came up to a very important conclusion: in more diversified regions there are more job opportunities and, hence, lower unemployment rates, since such regions are able to reduce the negative consequences of labour market shocks through a process of labour reallocation between different sectors. In other words, the more diversified is the region, the less damaging are the sectoral shocks affecting one or a small number of industries.

Simon and Nardinelli (1992) confirmed the hypothesis of a portfolio theory in the US labour market. They proved that with the growth of sectoral diversification, the influence of sectoral shifts on the production structure is reduced, but the probability of laid-off employees to find work in another industry due to the existence of Jacobs effects is higher. Mussida and Pastore (2015) found that sectoral changes lead to an increase in the unemployment level, while the existence of more specialized regions, according to the Marshall’s theory, partially neutralizes the negative consequences of specialization, expressed in greater exposure to external shocks. Mameli et al. (2008), as well as Paci and Usai (2008), confirmed the negative impact of specialization externalities and the positive effect of diversification on regional employment growth of Italy. Forni and Paba (2002) found out that both externalities from specialization and urbanization positively influence on the dynamics of employment.

The literature provides empirical evidence that confirms the presence of the Jacobs effect in a number of countries. However, Simon and Nardinelli (1992) showed that there were periods (for example, the beginning of the Great Depression), when unemployment was higher in more diversified regions. Since the period 2007-2016 is also not homogeneous, in particular, it contains 2008, when the financial crisis began, we conducted studies not only for the entire period under consideration, but also for shorter favorable and crisis periods. Beaudry and Schiffauerova (2009) conducted a meta-analysis of 67 studies and showed that, depending on the methodology and period of analysis, one of the effects predominates. Almost the same number of studies confirmed the positive influence of both effects on unemployment level. Similarly, De Groot et al. (2009) found the same results based on a meta-analysis of 31 articles on agglomerative externalities.

There are only a few studies devoted to the influence of the concentration and diversification on the main macroeconomic indicators of Russian regions, but the previous contributions left many gaps in the literature which we aim to bridge at least partly in our paper. Rastvortseva and Ternovsky (2016), calculated the Herfindahl-Hirschman index in three different ways (according to the number of employees, investments and the volume of production), and drew attention to the increase in concentration in the Russian economy from 1990 to 2013. Special features of localization effects in large Russian cities are discussed in Animitsa (2013), Rastvortseva, Manaeva (2015) and Bufetova, (2017). Maslikhina (2013) assessed "that the proceeding territorial concentration of economic activity will strengthen the spatial inequality". Demidova et al. (2018) estimated the employment models for Russia and found that in 2005-2016 in Russia the Marshallian effects prevailed.

The authors of most of these studies used linear dependencies between the concentration and diversification and indicators of labour market. However, some authors found evidence of a non-linear effect using quadratic functional form, Davidson and Mariev (2015), Vorobyov et al. (2014) revealed an inverted U-shape dependence between enterprise revenue (dependent variable) and the level of regional localization. At the same time the influence of diversification on enterprise revenue was positive.

Thus, it makes sense to use the nonmonotonic dependence of changes in unemployment from concentration or diversification for Russian data. The quadratic parameterization is only one of the possible nonlinear models, and, by its very nature, semiparametric modeling captures better nonlinear dependencies among variables. To the best of our information, only Basile et al. (2012) used semiparametric dependence in a paper based on Italian data and found some evidence of non-linear dependence while looking at the influence of regional specialization and diversification on unemployment growth in Italy in 2004-2008. The specificity of the Italian case is due to the fact that Italy is a country with high spatial heterogeneity of local labour markets, and there are significant differences in productivity between the North and South. The authors found, in particular, that at low specialization values, Jacobs effects dominate due to intersectoral mobility, but in regions with a higher level of specialization, the role of Marshallian externalities increases. Thus, in highly concentrated regions, the overall effect of spatial specialization on unemployment growth is not statistically significant. However, this article used only one time period, before the 2008 economic crisis. We are not aware of articles using the data for the period after the economic crisis of 2008, in which flexible nonparametric dependence would be used.

Researchers interested in studying the determinants of the main Russian macroeconomic indicators (such as economic growth or unemployment), noted that in the empirical models it is necessary to take into account the mutual influence of regions on each other using spatial lags. Vakulenko (2015), Semerikova (2015) discuss the bias in the estimates of the coefficients due to ignoring spatial effects. To avoid such bias, we included spatial lags in our model.

3. The main hypotheses and data for their verification

3.1. Main hypothesis

In this paper we analyze the data for 80 Russian regions over 10 years (2007-2016) provided by the Russian statistical agency Rosstat, www.gks.ru (see Appendix A). Data on some regions are missing (the Republic of Chechnya, the Republic of Crimea and Sevastopol). The Kaliningrad region is not included in the study because it has no common borders with other regions of Russia. During the reporting period, some regions underwent changes of an administrative-territorial nature. This altering of boundaries is taken into consideration, mitigated by means of a procedure of aggregation (see Appendix B).

Unemployment is one of the main labour market indicators, used in these cases. Following Overman and Puga (2012) and Basile et al., 2012), we use the logarithm of unemployment growth as dependent variable. The log difference of unemployment rates approximates the average percentage increase in unemployment over the period $[t_1, t_2]$ in region i:

$$Y_i^{[t_1,t_2]} = \frac{\ln U_i^{t_2} - \ln U_i^{t_1}}{t_2 - t_1} \quad (1)$$

Table 1 contains information about the unemployment rate and average unemployment growth over the period 2007-2016 (in logarithms). The maximum regional unemployment level exceeds the minimum level by more than twenty times.

Russia is a regionally heterogeneous country and, hence, overlapping effects of the Jacobsen and Marshall economies are possible. We identify and test the following hypotheses.

Hypothesis 1. *Unemployment growth is non-linearly dependent on the degree of concentration or diversification due to the possible overlapping of urbanization and localization.*

During periods of economic growth regions with a high degree of diversification have more favorable employment indicators due to Jacobs effects, as they spread among different industries in one region, and labour mobility contributes to a reduction in unemployment.

On the contrary, in crisis periods localization effects prevail due to the declining demand for products. In addition, the number of firms reduces due to the closure of small uncompetitive companies, which leads to firms understanding the need for mutual cooperation in order to minimize costs and to use joint innovations. Having studied the transition period in the Russian and Chinese economies (1990's), Galbraith et al. (2004) argue that the industries with the maximum level of concentration remained in a winning position and were less affected by the crisis, in terms of unemployment risk.

Simon and Nardinelli (Simon, 1988, Simon and Nardinelli, 1992), Tress (1938), McLaughlin (1930), Elhorst (2003), Ferragina and Pastore (2008) also confirmed the effects of urbanization and portfolio hypothesis. They concluded that with diversification increasing in a region, employment opportunities increase due to shifts between sectors, which leads, in turn, to lower levels of unemployment. However, the authors show that there are such crisis periods, when, in more diversified regions, the unemployment rate is higher.

Hypothesis 2. *The direction of influence of concentration or diversification on unemployment growth changes over time.*

We consider the following periods: 2007-2016 (overall period, rather heterogeneous), 2008-2010 (crisis period), 2010-2013 (recovery period) and 2013-2016 (slowdown in economic growth). Our short periods overlap, because some changes did not occur at the end of the year, but approximately in the middle. Since we use annual rather than quarterly or monthly data (which are not available for Russian regions), we had to include some years in two periods.

3.2. An empirical study of indices of spatial concentration and diversification

We follow different approaches to the measurement of indicators of industrial concentration and diversification and try to

Table 1
Descriptive statistics of unemployment and unemployment growth, 2007-2016

Variable	Period	Observations	Mean	Std. Dev.	Min	Max
unemployment	2007	80	7.433	5.726	1.3	47.3
unemployment	2008	80	8.067	6.214	1.6	55
unemployment	2009	80	9.711	5.635	3.5	52.9
unemployment	2010	80	8.725	5.386	2.3	49.7
unemployment	2011	80	7.876	5.257	1.9	48.8
unemployment	2012	80	6.773	5.248	1.1	47.7
unemployment	2013	80	6.663	4.896	1.5	43.7
unemployment	2014	80	6.253	3.691	1.4	29.8
unemployment	2015	80	6.67	3.695	2.1	30.5
unemployment	2016	80	6.723	3.677	1.6	30.2
Average log of unemployment growth	2007-2016	80	-0.00437	0.035491	-0.09682	0.115978

develop the ideas of (Neumann and Topel, 1991) Chiarini and Piselli (2000), and Robson (2009).

As we noted in the introduction, we used both concentration and diversification indices and two types of data for their calculation: firm and regional level data. This help us to test the robustness of our results.

Firms-level indices have been based on firm revenues of Russian companies for the period 2007-2016 in various regions. Firm revenues were obtained from the database "Ruslana" provided by Bureau Van Dijk. In total, we have used information about 12116 companies, operating in 24 manufacturing industries (code C) in accordance with the OKVED 2 classification: food, beverages, tobacco, textiles, clothing, leather and leather products, wood processing and manufacture of wood products and cork, paper and paper products, printing and copying media, coke and petroleum products, chemicals and chemical products, medicines, rubber and plastic products, other non-metallic non-mineral products, metallurgical, finished metal products (except machinery and equipment), computers, electronic and optical products, electrical equipment, machinery and equipment, motor vehicles, trailers and semi-trailers, other vehicles and equipment, furniture, other finished products, repair and installation of machinery and equipment. We focus on manufacturing industries, since the extraction of minerals and their primary processing are not of special interest because of the imperfect labour mobility in these industries, documented by Mikhailova (2017).

For regional-level indices we used data of value added by 15 types of economic activity, as listed on the Rosstat website (www.gks.ru), for each region, annually for the years 2007-2016: agriculture, forestry, fishing; mining and quarrying; manufacturing; production and distribution of electricity, gas and water; construction; wholesale and retail trade; repair of motor vehicles and motorcycles; accommodation and food service activities; information and communication; financial and insurance activities; real estate, rent and services activities; public administration and defense; compulsory social security; education; human health and social work activities; provision of other communal, social and personal services.

The most frequent diversification index used in the literature is the Herfindahl-Hirschman index. However, Vorobyov (Vorobyov, 2014) proposed other diversification index (formula 2):

$$ihh_i^t = \frac{\sum_{j=1}^{S}\left[\frac{pq_{ij}^t}{pq_i^t}\right]^{\frac{1}{S}} - 1}{\left(S^{1-\frac{1}{S}}\right) - 1}, \quad ihh_i^t \in [0; 1] \tag{2}$$

where i is a number of a region, j is a number of industry, S is a number of industries in the economy; pq_{ij}^t is a revenue in industry j in region i; pq_i^t is a revenue in all industries in region i.

The advantage of using P. Vorobyov's diversification index from our point of view is that

1) The higher the value of this index, the higher the level of diversification, which is more convenient for economic interpretation.

An increase in the very popular Herfindahl-Hirschman index, on the contrary, corresponds to a decrease in the level of diversification.

1) The Vorobyov's index takes values from the interval [0, 1]. It is easy to obtain from formula (2) that if the distribution of economic activity across S industries is the same, which corresponds to the greatest diversification of production, then the Vorobyov's index is equal to the limit value of 1.

The Herfindahl-Hirschman index in this case is 1 / S (which is 0, corresponding to the greatest diversification, only in the limit).

We used an analogue of formula (2) to calculate two diversification indices calculated from data for firms ($ihhmn_i^t$) and from the value added of the gross regional product ($ihhva_i^t$) in the following way:

$$ihhmn_i^t = \frac{\sum_{j=1}^{24}\left[\frac{pq_{ij}^t}{pq_i^t}\right]^{\frac{1}{24}} - 1}{\left(24^{1-\frac{1}{24}}\right) - 1}, \tag{3}$$

where i is number of a region; $j \in \{1, \ldots, 24\}$ is the number of manufacturing industries,

pq_{ij}^t is revenue of all firms in industry j in region i, pq_i^t is revenue in all firms in region i,
$ihhmn_i^t = 1$ is the equal distribution of firm turnover between industries (diversification),
$ihhmn_i^t = 0$ is the uneven distribution of firms' turnover in industries (lack of diversification);

$$ihhva_i^t = \frac{\sum_{j=1}^{15}\left[sh_{ij}^t\right]^{\frac{1}{15}} - 1}{\left(15^{1-\frac{1}{15}}\right) - 1}, \tag{4}$$

where i is the number of a region; $j \in \{1, \ldots, 15\}$ is the number of the type of economic activity,

t is a year,
sh_{ij}^t is the share of j-th type of economic activity in region i;
$ihhva_i^t = 1$ is the equal distribution of economic activity (diversification);
$ihhva_i^t = 0$ is the uneven distribution of economic activity (lack of diversification).

To test the robustness of our results we used also the second index of concentration from (Henderson, 2003). The Ellison-Glazer index of concentration is the sum by regions of the square deviation of the share of each region in the national revenue in industry j from its share in the national revenue.

Table 2
Minimum, maximum and average values of concentration and diversification indices, 2007-2016

Index	Minimum		Maximum		Average value	
	2007	2016	2007	2016	2007	2016
Ihhva	0.772	0.797	0.977	0.974	0.88	0.908
Ihhmn	0.084	0.082	0.978	0.973	0.715	0.705
Iegva	0.007	0.009	0.506	0.402	0.057	0.052
Iegmn	0.035	0.02	0.834	0.906	0.218	0.229

We calculated two concentration indices using data for firms ($iegmn_i^t$) and for value added of the gross regional product ($iegva_i^t$) in the following way:

$$iegmn_i^t = \sum_{j=1}^{24} \left(\frac{pq_{ij}^t}{pq_j^t} - \frac{pq_j^t}{pq^t} \right)^2, \tag{5}$$

pq_{ij}^t is revenue of all firms in industry j in region i,
pq_i^t is revenue in all firms in region i,
pq_j^t is a revenue in all firms in industry j,
pq^t is a revenue in all firms,
t is a year,

$$iegva_i^t = \sum_{j=1}^{15} (sh_{ij}^t - sh_j^t)^2, \tag{6}$$

where i is the number of a region; $j \in \{1, \ldots, 15\}$ is the number of the type of economic activity,

t is a year,
sh_{ij}^t is the share of j-th type of economic activity in region i,
sh_j^t is the share of j-th type of economic activity in Russia.

These indices take values from 0 to 2.

$iegmn_i^t = 2$ – specialization of the region on one industry is observed,
$iegmn_i^t = 0$ – the region does not specialize in one industry,
$iegva_i^t = 2$ – specialization of the region in one type of economic activity is observed,
$iegva_i^t = 0$ – the region does not specialize in one type of economic activity.

Table 2 presents descriptive statistics of each index for the first and last year.

Usually, the larger the number of industries, the higher the degree of disaggregation of the data and the higher is the degree of diversification. In this case we could expect to find a slightly higher degree of diversification in the firm-level index. However, in the firm-level case, we used only 24 manufacturing industries, and in the regional-level case for all types of activity, therefore, the opposite result is possible.

In fact, the diversification indices for the two types of data reported in Table 2 are not very different from each other, but the concentration indices for the firm-level data are slightly higher.

According to Table 2, both diversification indices take relatively large values, and both concentration indices are relatively small, which indicates that in Russia the level of diversification of production is quite high.

Indices calculated on revenues indicate an increase in the concentration of manufacturing industries 2007-2016, as the average value of the concentration index increased, and the average value of the diversification index, on the contrary, decreased in 2016 as compared to 2007. Indices calculated on the value added by economic activity, on the contrary, indicate an increase in diversification and a decrease in concentration. In addition, in 2016, there is a decrease in the spread between the minimum and maximum values for indices calculated by value added.

However, a spatial index reflecting the concentration or diversification in the region is only one of the possible variables that can affect the unemployment rate.

3.3. Other variables

Based on previous studies, we include the following control variables: GRP (gross regional product) per capita, calculated in the base prices of 2000, the share of urban population, the share of population with higher education, the coefficient of migration increase per 10000 people, the share of people below working age (below 16 years), the share of people above working age (55 years for women and 60 years for men), population density (number of persons per square kilometer), the Lilien index, initial unemployment level and the growth of weighted unemployment in neighboring regions (the spatial lag of the dependent variable).

Lilien (1982) found confirmation of the positive correlation over time between the aggregate unemployment rate and intersectoral variance in the growth of employment. He also created a statistical index for measuring changes in industries. In our paper this index is calculated as

$$lilien_i^t = \left(\sum_{j=1}^{15} \left[\frac{x_{ij}^t}{x_i^t} \right] \cdot (\Delta \ln x_{ij}^t - \Delta \ln x_i^t)^2 \right)^{1/2} \tag{7}$$

where x_{ij}^t is regional employment in the region i in type of economic activity $j \in \{1, \ldots, 15\}$,

x_i^t is total regional employment, Δ is a first difference operator.

The Lilien index reflects the dependence of labour demand on sectoral shifts in production. This index takes the value of 0, if no structural changes occurred during the period. The higher the value of the index, the faster the rate of structural change and more displacements in the labour market between sectors take place (Lilien, 1982). High values of the Lilien index lead to an increase in unemployment growth rates, especially for economically "weak" regions. Lehmann and Walsh (1999) proposed a possible explanation: when human capital can be exchanged, workers do not object to restructuring, which in turn increases unemployment, but provides a quick way out. High unemployment arises due to the mismatch of employer requirements and the abilities of the employee, and a low unemployment rate correlates with greater stability in the workplace. Samson (1985) was the first one to confirm Lilien's findings on Canadian data. Newell and Pastore (2006) also came to the same conclusions for the unemployment rate in Poland: high unemployment is a consequence of a mismatch between employer requirements and worker capabilities, and a low unemployment rate correlates with greater stability in the workplace. Krajnya'k and Sommer (2004) also found confirmation of the importance of the Lilien index in the Czech Republic in 1998-1999 at the time of economic restructuring. Abraham and Katz (1986) concluded that it is necessary to disentangle sectoral shifts and general market shocks, and noticed that the Lilien index truly describes sectoral shifts only if a measure of spatial diversity (the concentration or diversification index) is included among the regressors (Neumann and Topel, 1991).

A number of empirical works have proven that GRP negatively affects the unemployment rate, that is, Okun's law works. However, Elhorst (2003) also shows that this dependence of unemployment on GRP is not always observed. This suggests that the relationship can be nonlinear. This is why a nonparametric form of the relationship is more realistic: it allows for testing whether the relationship changes at different levels of the independent variable.

It is also difficult to predict the parametric form of the relationship between share of urban population and the unemployment growth. On the one hand, the unemployment level should increase with the rise in the share of urban population due to higher competition in the labour market, but with the growth of the already high values of the share of urban population, unemployment can decline as there are a lot of jobs in regions with a large number of urban population and job search takes less time due to more developed information mechanisms (lower frictional and mismatch unemployment) and increased density (Molho, 1995). Due to the ambiguous impact of this variable on unemployment growth, we expect to observe nonlinear dependence.

An increase in the share of people with higher education may have a two-way effect on the dynamics of unemployment. On the one hand, in regions with a low share of the population with high education, educated people find it difficult to find a job due to the lack of supply, which increases unemployment. But on the other hand, for regions with a high share of the population with higher education, its further growth stimulates a reduction in unemployment since in such regions the equilibrium state in the market is reached faster (Aragon et al., 2003). Thus, we expect a nonlinear U-shaped relationship between these two variables.

To control for possible endogeneity of the migration variable, we use the lag of migration net rate per 10,000 people. The correlation of this variable with the unemployment rate may also be nonlinear, namely bell-shaped. On the one hand, the influx of migrants occurs in favorable regions with low unemployment, where it is easy to find work. But, on the other hand, if there are too many such migrants, strong competition for jobs may arise.

The share of people below working age (up to 16 years) should be directly correlated with unemployment growth, because an increase in younger segment of the population means an increase in the extra labour force that will appear in the market and will be actively seeking jobs. With labor demand remaining constant, an increase in the youth labor force will make it more difficult to find a job. The increased competition will increase also unemployment. In addition, the unemployment risk is significantly higher for young people, due to their lack of work experience (Pastore, 2018), and a larger share of young people will typically be associated with increasing regional unemployment (Hofler and Murphy, 1989), and Elhorst, 1995).

In recent years, there has been an increase in economic activity among the elderly population, especially "young" pensioners[1]: they represent a sort of reservoir for employment growth. Moreover, there is a very low share of self-employed (about 2%) by European standards, which stimulates people above working age to continue working (Sonina, 2015). In addition, Russia is also experiencing the global increase in the number of years that people work (Sinyavskaya et al., 2017). However, at very high levels of unemployment, pensioners are likely to no longer be actively seeking work. Thus, there may be a non-linear relationship between the share of people above working age and the increase in unemployment level. Partridge and Rickman (1995) compared the effect of an increase in the share of the youth and elderly components of the labor force on unemployment and found that the impact of the former is much greater than the latter.

Population density is calculated as the number of people per square kilometer. Large and densely populated regions should have greater efficiency in finding work for their residents, therefore contributing to a lower unemployment rate (Elhorst, 2003). However, there is an opposite effect: population density reflects the convenience and greater attractiveness of large regions for life, which causes congestion effects, and, as a result, a higher level of unemployment (Niebuhr, 2003). In different time periods these effects can overlap, so we hypothesize the existence of a nonlinear relationship (u-shaped) (Basile et al., 2012).

Based on the work of Overman and Puga (2002), we included the logarithm of the unemployment rate of the region at the beginning of the period to assess whether the processes of beta convergence of regions in terms of unemployment take place.

One of the explanatory variables is the average increase in unemployment in neighboring regions (*wgrunempl*), which is calculated by multiplying the dependent variable by W (weighting matrix). In this paper we used a weighting binary matrix of dimension 80*80:

$$W = \begin{pmatrix} 0 & w_{12} & \dots & w_{1n} \\ w_{21} & 0 & \dots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \dots & 0 \end{pmatrix} \tag{8}$$

The elements of the weighting matrix are defined as follows: $w_{ij} = 1$, if the regions have common border and 0 otherwise. Then the elements of the weighting matrix were normalized in a row.

The effect of this variable on unemployment growth can be multidirectional. Basile et al. (2012) proved spatial dependence of unemployment growth in the Italian regions: the statistically significant coefficient suggest that neighboring regions showed a greater level of spatial "contamination" than regions located further apart. However, the impact may be the opposite: it is possible to reduce regional unemployment in response to the rise of unemployment in neighboring regions if the region attracts labour. Due to the possible existence of two opposite effects, we use a nonparametric dependence of unemployment growth in a region on weighted unemployment in neighboring regions.

Regarding the inclusion of control variables in our models. There is a trade - off between omitted variable bias when missing an essential variable and inefficiency of estimates when we add irrelevant factors in our models (while the estimates remain unbiased). As it will be shown in the next section, the dependence on each of the control variables was statistically significant at least in one estimated model. Therefore, for the convenience of comparing model estimates at different time intervals, we included a complete set of control variables in the models. To make sure that there is no multicollinearity problem, we calculated the condition number for matrix X, which includes indicators of interest ihhva, ihhmn, egva, egmn and all control variables. The condition numbers were equal to 10.093 for 2007 year, 10.374 for 2008 year, 11.763 for 2010 year, and 10.15 for 2013 year. In all cases, the condition numbers were less than 20, and according to Greene (2012), we have no multicollinearity problem. And at the same time we avoid omitted variable bias problems.

4. Model, estimation methodology and main results

4.1. Estimation methodology

As noted earlier in the modeling of unemployment growth, it is preferable to use a more flexible nonparametric functional form of dependence for all variables. We use an additive semi-parametric model, since the additivity property assumes that the effect of each explanatory variable in the model can be interpreted separately from other regressors, just as in linear multiple regressions. In addition, this model gives a graphical representation of the relationship between the dependent variable and the explanatory variables. The classical semiparametric additive model (AM) is:

$$Y_i = \beta_0 + \beta_1 X^*_{1i} + \beta_2 X^*_{2i} + \dots + f_1(X_{1i}) + f_2(X_{2i}) + \dots + \varepsilon_i \tag{9}$$

[1] Citizens permanently residing or working in the North and in areas similar to it, workers engaged in hard and harmful work, etc.

where $X_{1i}^{*}, X_{2i}^{*}, \ldots$ are strictly parametric components, $\beta_0,\ \beta_1,\ \beta_2, \ldots$ are the corresponding parameters, $f_1,\ f_2, \ldots$ are unknown smooth functions, ε is the vector of independent identically distributed errors (iid).

We use the methodology proposed by Wood (2006) and estimate $f_1,\ f_2, \ldots$ by means of penalized cubic regression spline. The selection of smoothing parameters was carried out using the generalized cross-validation method.

Skipping spatial autocorrelation (that reflects the average increase in unemployment in neighboring regions) can lead to the omission variable problem, incorrect estimates and conclusions. In order to control for spatial interaction, we include in the model the spatial lag of the dependent variable $WY_i^t = \sum_{j \neq i} w_{ij} Y_j^t$ (where w_{ij} are elements of the spatial weights matrix, which reflects the interaction between regions i and j).

The final spatial autoregressive additive models used in our paper is:

$$Y_i^{[t_1,t_2]} = \beta_0 + \beta_{X_1^*} X_{1i}^{t_1 *} + \beta_{X_2^*} X_{2i}^{t_1 *} + \ldots + f_{X_1}\left(X_{1i}^{t_1}\right) + f_{X_2}\left(X_{2i}^{t_1}\right) + \ldots + f_{WY}\left(WY_i^{[t_1,t_2]}\right) + \varepsilon_i^{t_1} \qquad (10)$$

where i is a region number, $[t_1,\ t_2] \in \{[2007,\ 2016], [2008, 2010], [2010,\ 2013], [2013,\ 2016]\}$ are periods of consideration, dependent variables are given by (1), all explanatory variables were listed above.

In fact, we use a lag structure for explanatory variables. Indeed, our explanatory variables are measured at time t1, and the dependent variable is the average change in unemployment rate (in logarithms) over a period of time $[t_1, t_2]$ (see formula 1). As we noted earlier, our whole time interval 2007-2016 is heterogeneous, so we considered more homogeneous and short periods with the starting points t1 = 2008 (beginning of the crisis), 2010 (beginning of the recovery period), 2013 (the beginning of a worsening the economic situation in Russia).

Since Y and its spatial lag WY may be interrelated, there is the problem of endogeneity. To control for this problem, we use the two-step approach proposed by Blundell and Powell (2003). This is an analog of the Durbin-Wu-Hausman algorithm in the linear case, used in the presence of endogenous regressors.

In the first step, the following auxiliary semiparametric regression is considered:

$$WY_i^{[t_1,t_2]} = \tilde{f}_{X_1^*}^{*}\left(X_{1i}^{t_1 *}\right) + \tilde{f}_{X_2^*}^{*}\left(X_{2i}^{t_1 *}\right) + \ldots + \tilde{f}_{X_1}\left(X_{1i}^{t_1}\right) + \tilde{f}_{X_2}\left(X_{2i}^{t_1}\right) + \ldots + \tilde{h}_{WX_1^*}^{*}\left(WX_{1i}^{t_1 *}\right) + \tilde{h}_{WX_2^*}^{*}\left(WX_{2i}^{t_1 *}\right) + \ldots + \tilde{h}_{WX_1}\left(WX_{1i}^{t_1}\right) + \tilde{h}_{WX_2}\left(WX_{2i}^{t_1}\right) + \ldots + \nu_i \qquad (11)$$

where explanatory variables $X_1^*,\ X_2^*, \ldots, X_1, X_2, \ldots$ were used as instruments for WY as well as their spatial lags $WX_1^*,\ WX_2^*, \ldots, WX_1, WX_2, \ldots$; and $\tilde{f}_{X_1^*}^{*}, \tilde{f}_{X_2^*}^{*}, \ldots, \tilde{f}_{X_1}, \tilde{f}_{X_2}, \ldots, \tilde{h}_{WX_1^*}^{*}, \tilde{h}_{WX_2^*}^{*}, \ldots, \tilde{h}_{WX_1}, \tilde{h}_{WX_2} \ldots$ are unknown smooth functions. We estimate them by means of a penalized cubic regression spline and back-fitting method of estimation[2] (Hastie and Tibshirani, 1990). ν_i are errors of regression.

The second step is to estimate an additive model of the following form:

$$Y_i^{[t_1,t_2]} = f_{X_1^*}\left(X_{1i}^{t_1 *}\right) + f_{X_2^*}\left(X_{2i}^{t_1 *}\right) + \ldots + f_{X_1}\left(X_{1i}^{t_1}\right) + f_{X_2}\left(X_{2i}^{t_1}\right) + \ldots + f_{WY}\left(WY_i^{[t_1,t_2]}\right) + f_{\nu}\left(\hat{\nu}_i\right) + \varepsilon_i^{t_1} \qquad (12)$$

This model includes the same explanatory variables as the original model and additionally a nonparametric function that depends on the model residuals obtained in the first step. Like in the first step we used back-fitting and penalized cubic regression splines for estimation of all unknown smooth functions. For each explanatory variable, we choose between linear and nonparametric dependence: the null hypothesis is that the dependence is linear, and the alternative hypothesis is that the dependence is nonparametric. In the absence of a significant difference, we opt for a linear form.

All calculations were performed in a statistical package R and RStudio with the help of the special package MGCV, which includes an estimate of the general additive model (gam). After conducting preliminary tests on the choice of linear or semiparametric dependence and the ANOVA test for each explanatory variable, we found that linear dependence applied for the variables share of people below working age (up to 16 years) and the Lilien index.

4.2. Testing the main hypotheses

Finally, we have chosen 16 models:

Model 1: $Y_i^{[2007,2016]} = \beta_0 + \beta_1 lilien_i^{2007} + \beta_2 youth_i^{2007} + \ldots + f_{ihhva}(ihhva_i^{2007}) + f_{grp}(grppercap_i^{2007}) + \ldots + f_{unempl}(unempl_i^{2007}) + f_{WY}(WY_i^{[2007,2016]}) + \varepsilon_i^{2007}$,

Model 2: $Y_i^{[2007,2016]} = \beta_0 + \beta_1 lilien_i^{2007} + \beta_2 youth_i^{2007} + \ldots + f_{ihhmn}(ihhmn_i^{2007}) + f_{grp}(grppercap_i^{2007}) + \ldots + f_{unempl}(unempl_i^{2007}) + f_{WY}(WY_i^{[2007,2016]}) + \varepsilon_i^{2007}$,

Model 3: $Y_i^{[2007,2016]} = \beta_0 + \beta_1 lilien_i^{2007} + \beta_2 youth_i^{2007} + \ldots + f_{egva}(egva_i^{2007}) + f_{grp}(grppercap_i^{2007}) + \ldots + f_{unempl}(unempl_i^{2007}) + f_{WY}(WY_i^{[2007,2016]}) + \varepsilon_i^{2007}$,

Model 4: $Y_i^{[2007,2016]} = \beta_0 + \beta_1 lilien_i^{2007} + \beta_2 youth_i^{2007} + \ldots + f_{egmn}(egmn_i^{2007}) + f_{grp}(grppercap_i^{2007}) + \ldots + f_{unempl}(unempl_i^{2007}) + f_{WY}(WY_i^{[2007,2016]}) + \varepsilon_i^{2007}$,

Model 5: $Y_i^{[2008,2010]} = \beta_0 + \beta_1 lilien_i^{2008} + \beta_2 youth_i^{2008} + \ldots + f_{ihhva}(ihhva_i^{2008}) + f_{grp}(grppercap_i^{2008}) + \ldots + f_{unempl}(unempl_i^{2008}) + f_{WY}(WY_i^{[2008,2010]}) + \varepsilon_i^{2008}$,

Model 6: $Y_i^{[2008,2010]} = \beta_0 + \beta_1 lilien_i^{2008} + \beta_2 youth_i^{2008} + \ldots + f_{ihhmn}(ihhmn_i^{2008}) + f_{grp}(grppercap_i^{2008}) + \ldots + f_{unempl}(unempl_i^{2008}) + f_{WY}(WY_i^{[2008,2010]}) + \varepsilon_i^{2008}$,

Model 7: $Y_i^{[2008,2010]} = \beta_0 + \beta_1 lilien_i^{2008} + \beta_2 youth_i^{2008} + \ldots + f_{egva}(egva_i^{2008}) + f_{grp}(grppercap_i^{2008}) + \ldots + f_{unempl}(unempl_i^{2008}) + f_{WY}(WY_i^{[2008,2010]}) + \varepsilon_i^{2008}$,

Model 8: $Y_i^{[2008,2010]} = \beta_0 + \beta_1 lilien_i^{2008} + \beta_2 youth_i^{2008} + \ldots + f_{egmn}(egmn_i^{2008}) + f_{grp}(grppercap_i^{2008}) + \ldots + f_{unempl}(unempl_i^{2008}) + f_{WY}(WY_i^{[2008,2010]}) + \varepsilon_i^{2008}$,

Model 9: $Y_i^{[2010,2013]} = \beta_0 + \beta_1 lilien_i^{2010} + \beta_2 youth_i^{2010} + \ldots + f_{ihhva}(ihhva_i^{2010}) + f_{grp}(grppercap_i^{2010}) + \ldots + f_{unempl}(unempl_i^{2010}) + f_{WY}(WY_i^{[2010,2013]}) + \varepsilon_i^{2010}$,

Model 10: $Y_i^{[2010,2013]} = \beta_0 + \beta_1 lilien_i^{2010} + \beta_2 youth_i^{2010} + \ldots + f_{ihhmn}(ihhmn_i^{2010}) + f_{grp}(grppercap_i^{2010}) + \ldots + f_{unempl}(unempl_i^{2010}) + f_{WY}(WY_i^{[2010,2013]}) + \varepsilon_i^{2010}$,

Model 11: $Y_i^{[2010,2013]} = \beta_0 + \beta_1 lilien_i^{2010} + \beta_2 youth_i^{2010} + \ldots + f_{egva}(egva_i^{2010}) + f_{grp}(grppercap_i^{2010}) + \ldots + f_{unempl}(unempl_i^{2010}) + f_{WY}(WY_i^{[2010,2013]}) + \varepsilon_i^{2010}$,

Model 12: $Y_i^{[2010,2013]} = \beta_0 + \beta_1 lilien_i^{2010} + \beta_2 youth_i^{2010} + \ldots + f_{egmn}(egmn_i^{2010}) + f_{grp}(grppercap_i^{2010}) + \ldots + f_{unempl}(unempl_i^{2010}) + f_{WY}(WY_i^{[2010,2013]}) + \varepsilon_i^{2010}$,

[2] Implemented in the R package mgcv.

Model 13: $Y_i^{[2013,2016]} = \beta_0 + \beta_1 lilien_i^{2013} + \beta_2 youth_i^{2013} + \ldots + f_{ihhva}(ihhva_i^{2013}) + f_{grp}(grppercap_i^{2013}) + \ldots + f_{unempl}(unempl_i^{2013}) + f_{WY}(WY_i^{[2013,2016]}) + \varepsilon_i^{2013}$,

Model 14: $Y_i^{[2013,2016]} = \beta_0 + \beta_1 lilien_i^{2013} + \beta_2 youth_i^{2013} + \ldots + f_{ihhmn}(ihhmn_i^{2013}) + f_{grp}(grppercap_i^{2013}) + \ldots + f_{unempl}(unempl_i^{2013}) + f_{WY}(WY_i^{[2013,2016]}) + \varepsilon_i^{2013}$,

Model 15: $Y_i^{[2013,2016]} = \beta_0 + \beta_1 lilien_i^{2013} + \beta_2 youth_i^{2013} + \ldots + f_{egva}(egva_i^{2013}) + f_{grp}(grppercap_i^{2013}) + \ldots + f_{unempl}(unempl_i^{2013}) + f_{WY}(WY_i^{[2013,2016]}) + \varepsilon_i^{2013}$,

Model 16: $Y_i^{[2013,2016]} = \beta_0 + \beta_1 lilien_i^{2013} + \beta_2 youth_i^{2013} + \ldots + f_{egmn}(egmn_i^{2013}) + f_{grp}(grppercap_i^{2013}) + \ldots + f_{unempl}(unempl_i^{2013}) + f_{WY}(WY_i^{[2013,2016]}) + \varepsilon_i^{2013}$.

The results of models (1)-(16), together with estimation and diagnostics tests for periods 2007-2016, 2008-2010, 2010-2013, and 2013-2016 are given in Appendices C, D.

Appendices E-H contain the fitted smooth functions $\hat{f}_{ihhmn}, \hat{f}_{ihhva}, \hat{f}_{iegmn}, \hat{f}_{iegva}$, if corresponding dependence is significant according to F-test. These graphs reflect the fitted one-dimensional smooth functions (solid lines), and the Bayesian 95% confidence intervals (grey areas), the details could be found in (Wood, 2004). On each graph, the vertical axis represents the level of the corresponding unemployment growth rates, and on the horizontal axis - the values of the explanatory variables.

The results obtained confirm our main hypotheses. The dependence of the dynamics of unemployment on the concentration or diversification in the general case is, indeed, nonlinear due to the overlap of the effects of urbanization and localization. In addition, the direction of their influence on unemployment growth depends on the specific time interval considered.

Only the dependence of the Ellison-Glaser index, calculated on the value added, was significant for the overall period (2007-2016). The dependence in this period is non-linear (see Appendix E, Fig. E.1): at low levels of concentration in the region, unemployment decreases with increasing concentration (thus, the localization effect predominates), but when the concentration exceeds a certain threshold value (ca 0.15), its further increase leads to a rise in unemployment (Jacobs externalities dominate).

The period 2007-2016, though, was quite heterogeneous, since Marshallian and Jacobs effects predominate in different years, so their overall effects overlap. That is why we lend special attention to the identification of the influence on unemployment growth in each period that reflects different economic situations.

In the crisis period (2008-2010), both diversification and concentration indices calculated on the basis of value added are statistically significant (see the Appendix F, Fig. F.1–3). Along with the diversification growth in the crisis period, the unemployment rate increases, indicating the predominance of Marshallian effects. Therefore, in 2008-2010, specialization effects prevail. These results are similar to the results for crisis periods in other countries (Simon and Nardinelli (Simon, 1988, Simon and Nardinelli, 1992), Tress (1938), McLaughlin (1930), Elhorst (2003), Ferragina and Pastore (2008).

The period 2010-2013 is one of “exit from the crisis” and of economic upsurge. In these years, the diversification and concentration indices, calculated on revenue, influenced unemployment growth (see Appendix G, Fig. G.1, Fig. G.2.). More specifically, with the increase in the diversification in the region, the unemployment rate is decreasing, and as concentration increases, unemployment grows, too (Jacobs effects are confirmed).

Finally, from 2013 to 2016, when the economic situation in the country began to deteriorate again (see Appendix H, Fig. H.1), a significant influence was confirmed for the diversification index calculated on revenue: an increase in diversification is associated with an upsurge in unemployment (the Marshallian effect predominates). On the level of diversification from 0.7 to 0.9, a small increase in the index leads to a decrease in unemployment (Jacobs effect for diversified regions), but an increase in the index value exceeding 0.9 rapidly increases unemployment. This is true for such regions as St. Petersburg, Yaroslavl Region, Leningrad Region, Moscow and Moscow Region, Krasnodar region. The dependence of the unemployment rate on the diversification index calculated by revenue is significant and non-linear.

Thus, during the period of economic recovery (2010-2013), people can find work in various industries and Jacobs effects prevail, and in difficult periods (such as 2008-2010 and 2013-2016) localization effects prevail.

Our assumptions about non-linear and non-monotonic dependences on the remaining variables have received only partial empirical confirmation. The effect of Lilien index on unemployment growth is statistically insignificant. it means that, in Russia, sectoral shifts in employment do not seem to influence unemployment growth. The positive influence of the share of the population below working age on the level of unemployment growth is empirically confirmed for 2007-2016 and 2008-2010 and could be attributed to the competition in the labour market.

We find positive dependence between the share of people above working age and unemployment change over the years 2007-2016 and 2013-2016 (see Appendix I), but when a certain share (ca 20%) is reached, its growth sharply slows down.

We also find evidence confirming our hypothesis about a U-shaped dependence of unemployment growth on population density in 2007-2016 and 2013-2016 (see Appendix J).

The dependence of unemployment change from spatial lag was positive and significant in 2007-2016, 2010-2013, which suggests that it is necessary to take into account the mutual influence of regions on each other (see Appendix K). Therefore, inclusion of spatial lags in the model was justified. Otherwise, we could have encountered the omitted variable bias problem.

5. Conclusions

It is impossible to draw unambiguous conclusions regarding which externalities predominate in the case of Russia due to the great regional heterogeneity and the presence of both urbanization and localization effects. In addition, the impact of the latter on unemployment growth is not constant over different time periods. During periods of economic growth (such as 2010-2013), people move between sectors and can easily find work, so that urbanization effects prevail, while in difficult periods (for example, 2008-2010 and 2013-2016), the localization effects dominate: the local agglomeration of firms from one industry creates a labour market with a limited set of skills that are on demand for a particular industry, and it is easier for people to find a job in industries of specialization.

Understanding the key differences among regions of the Russian Federation will allow the state to conduct a competent structured socio-economic policy that will help to eliminate the negative social and economic consequences from the high concentration in some regions. In crises time, the state should support enterprises whose specialization does not coincide with the main specialization of the region through tax benefits and special subsidies, and in periods of economic expansion, the government should develop the most promising sectors in each region. In addition, special attention should be paid to policies affecting youth to reduce their unemployment rate in regions with a high proportion of young people.

Appendix A. List of Russian regions

Number	Region	Number	Region
1	Belgorod region	41	Republic of Marii El
2	Bryansk region	42	Republic of Mordovia
3	Vladimir region	43	Republic of Tatarstan
4	Voronezh region	44	Republic of Udmurtia
5	Ivanovo region	45	Republic of Chuvashia
6	Kaluga region	46	Perm territory
7	Kostroma region	47	Kirov region
8	Kursk region	48	Nizhny Novgorod region
9	Lipetsk region	49	Orenburg region
10	Orel region	50	Penza region
11	Ryazan region	51	Samara region
12	Smolensk region	52	Saratov region
13	Tambov region	53	Ulyanovsk region
14	Tver region	54	Kurgan region
15	Tula region	55	Sverdlovsk region
16	Yaroslavl region	56	Tumen region
17	Moscow	57	Khanty-Mansi Autonomous Area - Yugra
18	Republic of Karelia	58	Yamal-Nenets autonomous region
19	Republic of Komi	59	Chelyabinsk region
20	Arkhangelsk region	60	Republic of Altay
21	Nenets Autonomous Okrug	61	Republic of Buryatia
22	Vologda region	62	Republic of Tyva
23	Leningrad region	63	Republic of Khakassia
24	Murmansk region	64	Altay Territory
25	Novgorod region	65	Zabaykalsky Territory
26	Pskov region	66	Krasnoyarsk Territory
27	Saint-Petersburg	67	Irkutsk region
28	Republic of Adygea	68	Kemerovo region
29	Republic of Kalmykia	69	Novosibirsk region
30	Krasnodar Territory	70	Omsk region
31	Astrakhan region	71	Tomsk region
32	Volgograd region	72	Republic of Sakha (Yakutia)
33	Rostov region	73	Kamchatka territory
34	Republic of Dagestan	74	Primorsky Territory
35	Republic of Ingushetia	75	Khabarovsk Territory
36	Republic of Kabardino-Balkaria	76	Amur region
37	Republic of Karachaevo-Cherkessia	77	Magadan region
38	Republic of Northen Osetia – Alania	78	Sakhalin region
39	Stavropol Territory	79	Jewish autonomous area
40	Republic of Bashkortostan	80	Chukotka Autonomous Okrug

Appendix B. United subjects of the Russian Federation

Data	Merging regions	Incorporated as
01.01.2007	Taymyr Autonomous Okrug Evenk Autonomous Okrug Krasnoyarsk territory	Krasnoyarsk Territory
01.07.2007	Kamchatka oblast Koryak Autonomous Okrug	Kamchatka territory
01.01.2008	Ust-Orda Buryat Autonomous Okrug Irkutsk region	Irkutsk region
01.03.2008	Chita region Aginsky Buryatsky Autonomous Okrug	Zabaykalsky Territory
01.07.2012	Moscow Moscow region	Moscow

Appendix C. Results of estimations, dependent variables are average unemployment growth (in log) over 2007-2016 and 2008-2010

Parametric terms (beta and p-values)	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Parametric terms (beta and p-values in brackets)								
time period	2007-2016	2007-2016	2007-2016	2007-2016	2008-2010	2008-2010	2008-2010	2008-2010
intercept	-0.227***	-0.230***	-0.235***	-0.256***	-0.280*	-0.252	-0.184	-0.323*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.090)	(0.137)	(0.249)	(0.068)
lilien	0.047	0.061	0.052	0.072	0.231	0.319	0.261	0.404
	(0.322)	(0.199)	(0.248)	(0.117)	(0.425)	(0.305)	(0.385)	(0.265)
young	0.013***	0.013***	0.013***	0.014***	0.019*	0.017*	0.013	0.021**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.056)	(0.096)	(0.173)	(0.048)

Nonparametric terms F test[a] and p-values in brackets		edf[b]		edf		edf		edf		edf		edf		edf		edf
f(ihhva)	1.225	1.949							4.861**	1.000						
	(0.273)								(0.032)							
f(ihhmn)			0.360	1.000							5.797**	1.000				
			(0.551)								(0.019)					
f(egva)					2.665*	1.935							9.244***	1.000		
					(0.070)								(0.003)			
f(egmn)							0.022	1.000							1.727	2.718
							(0.883)								(0.167)	
f(grppercap)	0.009	1.000	0.115	1.000	0.003	1.000	0.004	1.000	2.720	1.000	2.805	1.59	0.197	1.43	1.379	1.306
	(0.926)		(0.736)		(0.955)		(0.948)		(0.105)		(0.073*)		(0.71)		(0.229)	
f(urbanshare)	1.876	2.574	2.094	1.785	2.008	1.712	3.679*	1.538	0.719	1.157	0.709	1.599	0.756	1.64	0.448	1.0
	(0.149)		(0.132)		(0.153)		(0.053)		(0.518)		(0.518)		(0.476)		(0.506)	
f(higheduc)	2.960*	2.043	2.861*	1.995	2.489*	1.827	2.412*	1.750	3.054**	2.681	2.960**	3.179	3.141**	2.995	2.893**	292
	(0.051)		(0.053)		(0.084)		(0.095)		(0.031)		(0.032)		(0.026)		(0.0416)	
f(migr)	0.924	2.838	0.008	1.000	0.024	1.000	0.200	1.000	3.287**	3.808	3.78***	3.832	3.172**	3.797	1.701	3.54
	(0.524)		(0.931)		(0.877)		(0.656)		(0.016)		(0.007)		(0.018)		(0.157)	
f(old)	12.829***	2.514	14.143***	2.606	13.299***	2.648	15.157***	2.585	0.305	1.421	0.465	1.52	0.187	1.0	1.459	1.73
	(0.000)		(0.000)		(0.000)		(0.000)		(0.601)		(0.539)		(0.667)		(0.236)	
f(density)	12.388***	2.027	11.219***	2.019	14.519***	2.035	13.551***	2.008	2.829**	3.929	0.021	1.0	0.026	1.11	0.103	1.0
	(0.000)		(0.000)		(0.000)		(0.000)		(0.031)		(0.88)		(0.86)		(0.74)	
f(unempl)	28.557***	3.650	28.980***	3.618	34.057***	3.836	32.961***	3.851	13.124***	3.429	13.367***	2.886	11.87***	3.394	11.618***	3.275
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	
f(WY)	2.492*	1.117	2.084*	1.394	3.870**	1.879	3.683**	1.583	2.588*	1.107	0.038	1.0	0.766	1.0	0.496	1.0
	(0.094)		(0.095)		(0.026)		(0.024)		(0.090)		(0.94)		(0.38)		(0.48)	
R2 adj.	0.724		0.726		0.744		0.743		0.640		0.6		0.616		0.587	
GVC score	0.000		0.000		0.000		0.000		0.007		0.007		0.007		0.000	
Number of observations	80		80		80		80		80		80		80		80	

[a] F tests are used to test the overall significance of smooth terms.
[b] Effective degrees of freedom (edf) Significance levels: *** - less or equal 0.01, ** - less or equal 0.05, * - less or equal 0.1.

Appendix D. Results of estimations, dependent variables are average unemployment growth (in log) over 2010-2013 and 2013-2016

Parametric terms (beta and p-values)	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14	Model 15	Model 16
Parametric terms (beta and p-values in brackets)								
time period	2010-2013	2010-2013	2010-2013	2010-2013	2013-2016	2013-2016	2013-2016	2013-2016
intercept	-0.261	-0.279	-0.234	-0.150	-0.318	-0.314	-0.358	-0.306
	(0.012)	(0.005)	(0.015)	(0.139)	(0.000)	(0.000)	(0.000)	(0.001)
lilien	-0.172	-0.088	-0.124	-0.219	0.107	0.088	0.027	0.081
	(0.354)	(0.627)	(0.485)	(0.225)	(0.184)	(0.291)	(0.727)	(0.344)
young	0.010	0.011	0.008	0.004	0.017	0.017	0.020	0.017
	(0.080)	(0.052)	(0.121)	(0.496)	(0.000)	(0.000)	(0.000)	(0.001)

Nonparametric terms F test[a] and p-values in brackets		edf[b]		edf		edf		edf		edf		edf		edf		edf
f(ihhva)	0.476	1.000							1.778	1.000						
	(0.493)								(0.188)							
f(ihhmn)			4.529**	1.000							4.662***	3.928				
			(0.037)								(0.003)					
f(egva)					0.955	1.000							1.439	1.647		
					(0.333)								(0.175)			
f(egmn)							2.584*	2.870							1.600	1.000
							(0.060)								(0.211)	
f(grppercap)	0.027	1.000	0.532	1.517	0.194	1	1.119	2.399	4.516***	3.862	2.074*	3.435	0.625	1	1.454	3.246
	(0.869)		(0.499)		(0.661)		(0.321)		(0.003)		(0.096)		(0.432)		(0.224)	
f(urbanshare)	1.889	2.744	1.43	2.368	2.065	2.868	3.214**	2.871	9.396***	1	3.941*	1	3.103*	1.798	7.297***	1
	(0.129)		(0.21)		(0.098)		(0.022)		(0.003)		(0.052)		(0.072)		(0.009)	
f(higheduc)	1.69	1.0	0.264	1.0	0.195	1	0.586	1.04	2.213	1	0.692	1	0.11	1	1.971	1
	(0.198)		(0.609)		(0.66)		(0.485)		(0.142)		(0.409)		(0.741)		(0.166)	
f(migr)	2.389	1.000	3.787*	1.000	1.229	1	3.982*	1	0.003	1	0.607	1	1.558	3.662	0.098	1.25
	(0.127)		(0.056)		(0.272)		(0.050)		(0.955)		(0.439)		(0.161)		(0.757)	
f(old)	0.955	2.213	2.124*	2.389	0.922	1.717	0.405	1.502	8.995***	3.353	8.373***	2.736	9.035***	2.172	7.212***	2.756
	(0.302)		(0.081)		(0.374)		(0.546)		(0.000)		(0.000)		(0.000)		(0.000)	
f(density)	3.101**	2.903	2.995**	2.961	3.245**	2.932	4.213***	3.377	0.985	1.139	7.352***	2.775	4.373***	3.074	3.197**	3.047
	(0.024)		(0.025)		(0.019)		(0.003)		(0.357)		(0.000)		(0.006)		(0.028)	
f(unempl)	4.586**	2.441	3.693**	2.185	3.316**	0.945	3.327**	0.929	33.882***	0	8.683***	0	9.149***	0.983	7.365***	0.902
	(0.011)		(0.025)		(0.039)		(0.047)		(0.000)		(0.000)		(0.000)		(0.000)	
f(WY)	10.844***	1.0	7.48***	1.174	7.257***	1.76	5.268***	1.859	4.835***	3.845	0.082	1	4.72**	1	1.496	3.697
	(0.001)		(0.008)		(0.001)		(0.006)		(0.0017)		(0.775)		(0.034)		(0.192)	
R2 adj.	0.487		0.515		0.537		0.556		0.502		0.554		0.501		0.481	
GVC score	0.002		0.0019		0.0019		0. 0019		0.0016		0.0016		0.001		0.001	
Number of observations	80		80		80		80		80		80		80		80	

[a] F tests are used to test the overall significance of smooth terms.
[b] Effective degrees of freedom (edf) Significance levels: *** - less or equal 0.01, ** - less or equal 0.05, * - less or equal 0.1.

Appendix E. Partial effects of the index of concentration, 2007-2016

Fig. E.1

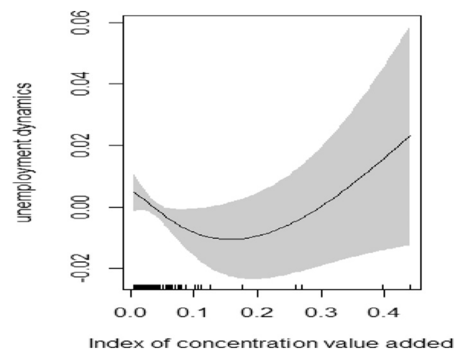


Fig. E.1. Index of concentration value added, model 3

.

Appendix F. Partial effects of the indices of concentration and diversification, 2008-2010

Figs. F.1.–F.3

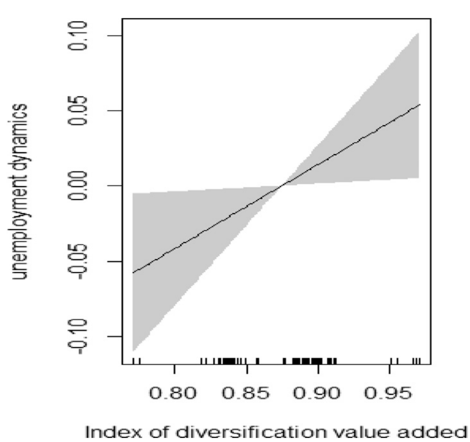


Fig. F.1. Index of diversification value added, model 5

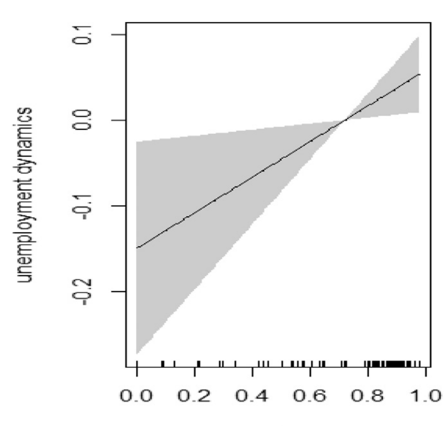


Fig. F.2. Index of diversification revenue, model 6

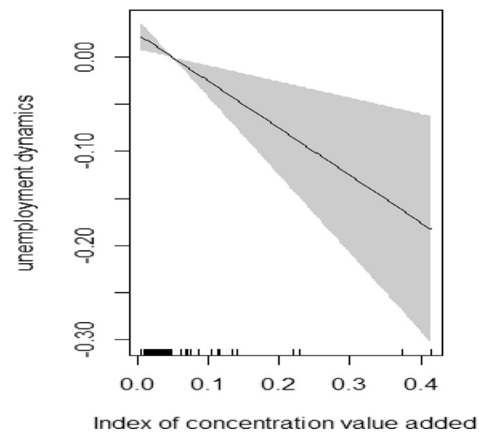


Fig. F.3. Index of concentration value added, model 7

.

Appendix G. Partial effects of indices of concentration and diversification, 2010-2013

Figs. G.1.–G.2

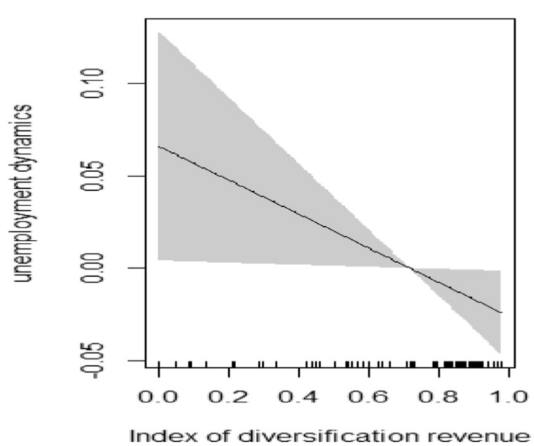


Fig. G.1. Index of diversification revenue, model 10

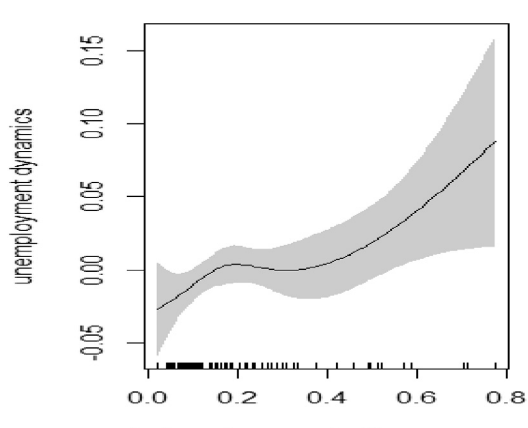


Fig. G.2. Index of concentration revenue, model 12

Appendix H. Partial effects of the index of diversification, 2013-2016

Fig. H.1.

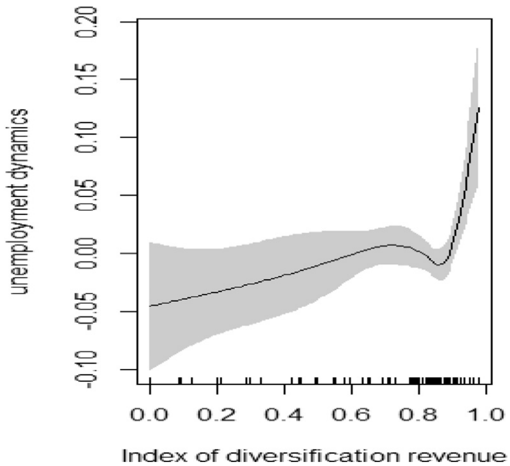
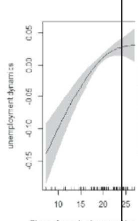
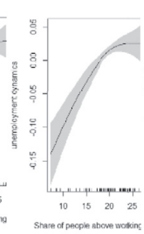
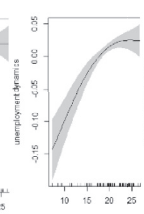
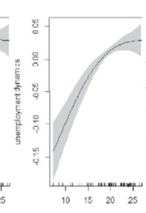
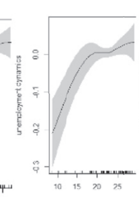
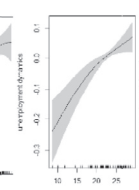
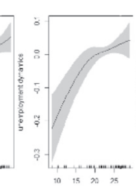
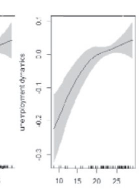
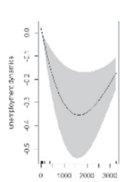
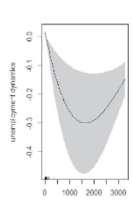
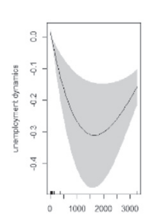
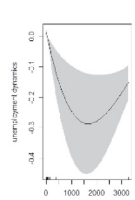
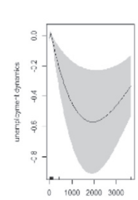
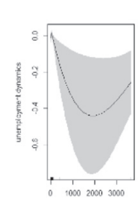
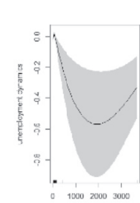


Fig. H.1. Index of diversification revenue, model 14

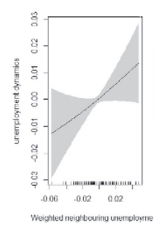
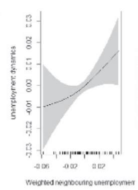
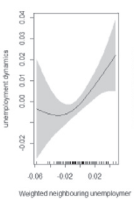
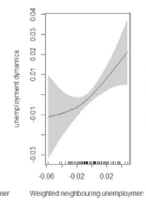
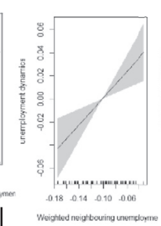
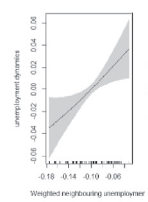
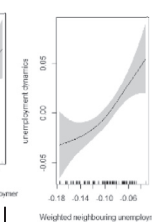
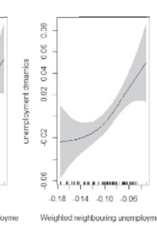
Appendix I. Partial effects of the variable "the share of people above working age"

Model	Model1	Model2	Model3	Model4	Model 13	Model 14	Model 15	Model 16
Time period	2007-2016	2007-2016	2007-2016	2007-2016	2013-2016	2013-2016	2013-2016	2013-2016
Functional form								

Appendix J. Partial effects of the variable "population density"

Model	Model1	Model2	Model3	Model4	Model1 4	Model1 5	Model 16
Time period	2007-2016	2007-2016	2007-2016	2007-2016	2013-2016	2013-2016	2013-2016
Functional form							

Appendix K. Partial effects of the spatial lag

Model	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Time period	2007-2016	2007-2016	2007-2016	2007-2016	2010-2013	2010-2013	2010-2013	2010-2013
Functional form								

CRediT authorship contribution statement

Olga Demidova: Conceptualization, Formal analysis, Methodology, Software, Supervision, Writing - original draft, Writing - review & editing. **Alena Kolyagina:** Data curation, Formal analysis, Software, Writing - original draft. **Francesco Pastore:** Conceptualization, Methodology, Software, Supervision, Writing - original draft, Writing - review & editing.

References

Abraham, K.G., Katz, L.F., 1986. Cyclical unemployment: sectoral shifts or aggregate disturbances? J. Political Econ. 94 (3), 507–522 1.
Animitsa, E.G., 2013. Largest Russian cities in the context of global urbanization processes. Ars. Administrandi 1, 82–96 (in Russian).
Aragon, S.R., 2003. Creating social presence in online environments. In: Aragon, Steven R. (Ed.). Facilitating Learning in Online Environments, New Directions for Adult and Continuing Education, 100. Jossey-Bass, San Francisco.
Basile, R., Girardi, A., Mantuano, M., Pastore, F., 2012. Sectoral shift, diversification and regional unemployment: evidence from local labour systems. Empirica 39, 525–544.
Beaudry, C., Schiffauerova, A., 2009. Who'sright,MarshallorJacobs?The localization versusurbanizationdebate. Res. Policy 38, 318–337.
Blanchard, O.J., Kremer, M., 1997. Disorganisation. Q. J. Econ. 112 (4), 1091–1126.
Blundell, R., Powell, J., 2003. Endogeneity in nonparametric and semiparametric regression models. In: Advances in Economics and Econometrics., 36. Cambridge University Press, Cambridge, pp. 312–357.
Bufetova, A.N., 2017. Trends in the concentration of economic activity and disparities in Russia's spatial development. Reg. Res. Russia 7 (2), 120–126.
Chiarini, B., Piselli, P., 2000. Unemployment, wage pressure and sectoral shifts: permanent and temporary consequences of intersectoral shifts. J. Policy Model. 22 (7), 777–799.
Davidson, N., Mariev, O., 2015. The impact of spatial concentration on enterprise performance. Econ. Reg. (4) 95–105.
De Groot, H., Poot, J., Smit, M.J., 2009. Agglomeration externalities, innovation and regional growth: theoretical perspectives and meta-analysis. In: Capello, R., Nijkamp, P. (Eds.), Handbook of Regional Growth and Development Theories.. Edward Elgar, Cheltenham, pp. 256–281.
Demidova, O., Daddi, P., Medvedeva, E., Signorelli, M., 2018. Modeling the employment rate in Russia: a spatial-econometric approach. Economy of Region 14 (4), 1383–1398.
Duranton, G., Puga, D., 2014. Micro-foundations of urban agglomeration economies. In: Henderson, J.V., Thisse, J.F. (Eds.). Handbook of Regional and Urban Economics, Vol. 4.
Elhorst, J.P., 2003. The mystery of regional unemployment differentials: theoretical and empirical explanations. J. Econ. Surv. 17 (5), 709–748.
Elhorst, J.P., 1995. Unemployment disparities between regions in the European Union, Convergence and Divergence Among European Unions. Pion, London, pp. 190–200.
Ferragina, A.M., Pastore, F., 2008. Mind the gap: unemployment in the New EU regions. J. Econ. Surv. 22 (1), 73–113.
Forni, M., Paba, S., 2002. Spillovers and the growth of local industries. J. Ind. Econ. 50, 151–171.
Galbraith, J.K, Krytynskaia, L., Wang, Q., 2004. The experience of rising inequality in Russia and China during the transition. Eur. J. Compar. Econ. Castellanza 1 (1), 87–106.
Gimpelson, V.E., Kapeliushnikov, R., 2016. Polarization or upgrading? Evolution of employment in transitional Russia. Russia. J. Econ. 2 (2), 192–218.
Greene, W., 2012. Econometric Analysis. Prentice Hall, Upper Saddle River, p. 90.
Hastie, T., Tibshirani, R., 1990. Generalized Additive Models. Chapman Hall & CRC. Monographs on Statistics & Applied Probability, 1. Chapman and Hall/CRC.
Henderson, J.V., 2003. Marshall's scale economies. J. Urban Econ. 53, 1–28.
Hofler, R., Murphy, K., 1989. Using a composed error model to estimate the frictional and excess–supply componets of unemployment. J. Reg. Sci. 29, 213–228.
Krajnya'k K., Sommer M. (2004), Czech Republic, IMF country report, August, 4/265.
Lehmann, H., Walsh, P.P., 1999. Gradual restructuring and structural unemployment in Poland: a legacy of central planning. LICOS DP, Centre for Transition Economics, 78. Catholic University of Leuven.
Lilien, D.M., 1982. Sectoral shifts and cyclical unemployment. J. Political Econ. Univ. Chicago Press 90 (4), 777–793.
McLaughlin, G.E., 1930. Industrial diversification in American cities. Q. J. Econ. 45 (1), 131–149.
Mameli, F., Faggian, A., McCann, P., 2008. Employment growth in Italian local labour systems: issues of model specification and sectoral aggregation. Spat. Econ. Anal. 3 (3) 343-259.
Marshall, A., 1920. Principles of Economics. Macmillan, London.
Maslikhina, V.Yu., 2013. Spatial heterogeneity of economic development of regional systems in Russia, Vestnik of the Volga State Technological University. Ser. Econ. Manag. 1 (17), 5–16 (in Russian).
Mikhailova T.N. (2017), Joint localization of Russian industries, Moscow, 74. (in Russian).
Molho, I., 1995. Migrant inertia, accessibility and local unemployment. Economica 62 (245), 123–132.
Mussida, C., Pastore, F., 2015. Geographical Labour Market Imbalances, AIEL Series in Labour Economics. Springer, Berlin and Heidelberg, p. 370.
Newell, A., Pastore, F., 2006. Regional unemployment and industrial restructuring in Poland. East. Eur. Econ. 44 (3), 5–28.
Neumann, G.R., Topel, R.H., 1991. Employment risk, diversification, and unemployment. Q. J. Econ. 106 (4), 1341–1365.
Overman, H., Puga, D., 2002. Unemployment clusters across European countries and regions. Econ. Policy 34, 117–147.
Paci, R., Usai, S., 2008. Agglomeration economies, spatial dependence and local industry growth. Revue d'économie Ind. 87–109.
Parr, J.B., 2012. Spatial-structure differences between urban and regional systems. Ann. Reg. Sci. 49 (2), 293–303.
Partridge, M., Rickman, D., 1995. Differences in state unemployment rates. The role of labour and product market structural shifts. South. Econ. J. 62, 89–106.
Pastore, F., 2018. Why is youth unemployment so high and different across countries? IZA World Labor 420. doi:10.15185/izawol.420.
Rastvortseva, S.N., Kuga, Ya.T., 2012. Regional specialization and geographical concentration of industries in Russia. Sci. Bull. BelSU 13 (132), 37–47 23/1(in Russian).
Rastvortseva, S., Ternovskii, D.S., 2016. Drivers of concentration of economic activity in Russia's regions // economic and social changes: facts, trends. Forecast 2 (44), 153–170.
Rastvortseva, S.N., Manaeva, I.V., 2015. Agglomeration processes in Russian cities. City Manag. Theory Pract. 3, 40–46 (in Russian).

Robson, M., 2009. Structural change, specialization and regional labour market performance: evidence for the UK, applied economics. Taylor & Francis J. 41 (3), 275–293.
Samson, L., 1985. A study of the impact of sectoral shifts on aggregate unemployment in Canada. Can. J. Econ. 18, 518–530.
Semerikova, E., Demidova, O., 2015. Analysis of regional unemployment in Russia and Germany: spatial-econometric approach. Spatial Econ. 2, 64–85 (in Russian).
Simon Curtis, J., Nardinelli, C., 1992. Does industrial diversity always reduce unemployment? Evidence from the great depression and after. Econ. Inq. 30, 384–397.
Simon, C., 1988. Frictional unemployment and the role of industrial diversity. Q. J. Econ. 103 (4), 715–728.
Sinyavskaya, O., Ermolina, A., Lyubushina, E., 2017. Employment after retirement in Russia: the role of pensions – active longevity or forced necessity? Matherials of the XVIII April International Scientific Conference of the Higher School of Economics (in Russian).
Sonina, Yu., Kolosnitsyna, M., 2015. Pensioners in the Russian labor market: trends in the economic activity of people of retirement age. (in Russian). Demographic Review 2 (2), 37–53.
Tress, R.C., 1938. Unemployment and the diversification of industry 1. Manch. Sch. 9 (2), 140–152.
Vakulenko, E., 2015. Analysis of the relationship between regional labour markets in Russia using Okun's model. Applied Econometrics. 40 (4), 28–48 (in Russian).
Vorobyov, P.V., Davidson, N.B., Kislyak, N.V., Kuznetsov, P.D., 2014. Industrial diversity and concentration in the Russian cities as factors of economic efficiency, Vestnik of the UrFU. Ser. Econ. Manag. 6, 4–18 (in Russian).
Voskoboynikov, I., 2017. Structural change, expanding informality and labour productivity growth in Russia/ NRU higher school of economics. Ser. WP BRP "Econ./EC" No. 168.
Wood, S.N., 2006. Generalized Additive Models. An Introduction with R. Chapman & Hall/CRC, Boca Raton, p. 397.
Wood, S.N., 2004. Stable and efficient multiple smoothing parameter estimation for generalized additive models. J. Am. Stat. Asso. 99 (467), 673–686.