Analysis of the regional disparities in Russia through STATIS methodology

Floro E. Caroleo*, Gianluigi Coppola**, Elena Semerikova***

Abstract

Russia is one of the most important nation of the world, with 145 million inhabitants and 84 regions. The process of structural change in the economic transition has been influenced by its size, distance to markets, climate, natural resource endowments and allocation of industries producing strong regional differentials. In this paper we analyse the differences among regions and their dynamics through the period between 2007 to 2013. For this purpose, we apply a dynamic multivariate method, named STATIS in order to individuate the main socio-economic characteristics of the regions, to find homogeneous clusters, and to examine their temporal dynamics. It can therefore be used to verify whether structural features favour the formation of clusters of regions and whether these display a tendency to converge either to a single structure or instead to a multiplicity of socio-economic structures.

JEL Classification: P25, R12, O18, C38

Key words: Russia regions, dynamic multivariate factorial analysis, STATIS, three way matrices.

Introduction^

The last decades have been characterized by a marked change of the economic and political geography of the European continent. As regards the Central and Eastern European countries, after the economic transition from a planned to a market economy leading to deep economic and institutional reforms, they have been exposed to other potential sources of structural change such

* Floro Ernesto Caroleo, Business and Economics Department (DISAE) and Intitute of Labour Economics (IZA) - Email ernesto.caroleo@uniparthenope.it, Correspondent author .

** Gianluigi Coppola, Department of Economics and Statistics (DISES), Centro di Economia del Lavoro e di Politica Economica (CELPE) – Email glcoppola@unisa.it .

*** Elena Semerikova, National Research University Higher School of Economics, Department of Applied Economics and Humboldt University, School of Business and Economics – Email semerlen@gmail.com.

^ the authors want to thank the participants to the Second World Congress of Comparative Economics, HSE University (St. Petersburg 14-17 June 2017) and to the XXXVIII Conferenza scientifica annuale AISRe (Cagliari 20-22 sept. 2017) for their very useful comments. Moreover, the authors are very grateful to Alessandro De Iudicibus and Luca Pennacchio for their help in data processing.

as the accession to the UE of most of them, the opening to world trade and more recently the consequences of the global financial crisis (Marelli and Signorelli 2010).

Looking at what happened in the old Europe after the creation of the European Union, an increasing interest has focused in particular the analysis of the causes of the socio-economic differences among the regions. In fact, despite the belief that a broader area of free trade would be a necessary and sufficient condition for economic welfare to spread uniformly among countries, the reality has proved that disparities among regions have been significantly greater than those among countries. The underlying hypothesis is that the structural change has strong spatially asymmetric impact on local labour markets. As a matter of fact, the excessive rigidity and the scant mobility (Barro and Sala-i-Martin 1992; Blanchard and Katz 1992; Decréssin and Fatas 1995; Obstfeld and Peri 1998; Boldrin and Canova 2001) of the labour factor, as well as the sectorial and institutional characteristics (Marelli et al. 2012), are judged to be the main causes of the persistence or of the increasing of the disparities among regions. Therefore, the literature that has analysed the convergence-divergence processes has paid increasing attention to the institutional mechanisms regulating the labour market, as well as to the characteristics of the labour supply and demand and their dependence on spatial factors (Erlhost 2000; Niebhur 2002). A series of regional factors connected with the labour market, often complementary but sometimes concomitant in creating divergence/convergence among regions have been consequently studied, as for example: the endowment of factors and 'fundamentals'; the structure of the labour market - natural growth and the age composition of the population, the composition of the labour force; migratory phenomena and commuting; the employment level, the gross regional product, the market potentials, the sectoral mix; density and urbanization; economic and social barriers, schooling levels; the institutional structure that regulates the goods and labour markets, or the composition of wages (Amendola et al. 2006)

As said, most Central and Eastern countries have experienced similar processes of structural change. In particular, the specific nature of economic transition, causing massive industrial and institutional restructuring, has exacerbated the dramatic and persistent labour market consequences and the spatially asymmetric impact on local labour markets (a broad review of the literature can be found in Pastore 2012, Caroleo and Pastore 2009, Caroleo and Pastore 2010). Moreover, the economic crisis, which began in most European countries in mid-2008, has had severe effects on EU and Eastern Europe labour markets. Although no country has been able to escape the crisis, the extent of output loss and the number of jobs lost, as well as the resulting rise in unemployment, vary considerably among countries and regions. Overall, the recession has not affected all workers in the same way: the low-educated young people and women have turned out to be the most vulnerable groups to the crisis (Marelli et al. 2012).

The aim of this paper is to examine the regional imbalance of Russia applying a dynamic multivariate factorial analysis method (the STATIS method) which, we believe, lends itself well to verification of most of the phenomena just described. The STATIS method, in fact, enables the Russian regions to be 'read' on the basis of factors that sum up their main socio-economic characteristics, to group them into homogeneous clusters, and to examine their temporal dynamics. It can therefore be used to estimate whether structural features favour the formation of clusters of regions and whether these display a tendency to converge either to a single structure or instead to a multiplicity of socio-economic structures. On this basis, it is then possible to investigate a number of issues: among them, what criteria could be used in defining regional or national policies or what institutional arrangement could better favour the development of a region.

In the first section the main characteristics of the Russian regions are described while the second section provides a brief description of the STATIS method and of the data set utilized. In the third, the method is applied to the Russia regions and the analysis is conducted of the characteristics of the main clusters of regions and of their dynamics over time. The concluding section provides a summary of the results.

1) The main characteristics of the Russian regions

Along with old members of European Union, regional imbalances have become prevalent also in other countries. Russia is a bright example of a country with a very diversified economic development in different territories. This diversity in economic development is mainly caused by spatial inequality in natural resources, infrastructure, climate and others. At the same time natural resources and geographical location are not the only reasons for regional development. Along with these factors regional policy might play a huge role. It is crucial for the economy of a region if the natural resources are used effectively or if the lack of natural resources is covered with the inner reserves. All these aspects are reflected in social and economic indicators and hence, cause economic and social regional differentiation (Ickes and Ofer, 2006).

Due to these differences different regions are characterized by different levels of economic development as well as by different regional specializations. Some regions focus on mining (for instance regions in Siberia, Ural), others on agriculture (Povolzhie, South regions), some regions are industrial (Ural and others).

Leading positions in regional economic development have Moscow, Saint-Petersburg, Khanty-Mansi Autonomous Area - Yugra, Moscow Region and the Republic of Tatarstan. Due to strong fundamental economic background these regions retain leading positions for most indicators of economic and social development. Other developed regions are Yamalo-Nenets Autonomous District, Sakhalin region, the Tyumen region, Sverdlovskaya region and the Republic of Bashkortostan.

So, looking at the list of the most developed regions one can notice the structure of the Russian economy: Moscow and Saint-Petersburg, financial and intellectual centers of the Russian Federation, and several industrial regions and regions with the big amount of gas and oil. All together these 10 above mentioned regions earn more than 70% of all oil mined in the country and 90% of gas. Least developed regions are Sevastopol, Republic of Adygea, Karachay-Cherkess Republic, Kabardino-Balkarian Republic, Republic of North Ossetia-Alania, Republic of Kalmykia, Republic of Altai, Republic Ingushetia, Jewish Autonomous Oblast, Republic of Tyva.

An important indicator of economic development of a region is investment per capita. The highest amount of investment per capita are obtained by Nenetskiy autonomous okrug and Yamalonenetskiy avtonomniy okrug. Due to the crisis of 2008-2009 investments declined sharply. The sharpest decline of investment was in Sakhalin and Nenetskiy autonomous okrug. There was a hug drop in investments in Kaliningrad, Samarskaya region, Moscow region, Moscow and Saint-Petersburg. That time some less developed regions in the Central and South part of Russia received transfers from the federal budget for houses construction purposes. However, 25 regions experienced higher level of investment than before the crisis. Highest growth of investment was in Chechnya, Ingushetiya, Primorskiy region and Krasnodarskiy region. There was also an investment growth in regions of new oil fields (Krasnoyarsk region, Komi Republic) and places of an oil tube on construction (Chabarovskiy region). New investments were made in Komi Republic, Yakutiya, Amurskaya regions and others.

Purchasing power of a population is the highest in Nenetsky avtonomniy okrug, Yamalo-Nenetskiy avtonomniy okrug and Moskov. The lowest purchasing power is in Republic Kalmikiya, Republic Ingushetia, Republic Tyva, Altai Republic, Krym Republic, Karachaevo-Cherkesskaya Republic. Concerning characteristics of the labour market Moscow and Saint-Peterburg have the lowest unemployment rate, whereas Ingushetiya Republic has the highest unemployment rate. Hence, Russian regions experience huge disparities in standard of living across regions.

These huge disparities in Russia are usually studied with respect to some particular indicators such as GDP and labor market indicators. Analysis of Russian regional data with respect to the spatial behavior is performed by Solanko (2003), Lugovoy (2007), Kolomak (2010), Ivanova (2015), Perret (2016). Demidova (2015) studies spatial effects of the main macroeconomic indicators of the eastern and western regions of Russia. Semerikova and Demidova (2015) analyze spatial disparities between regional unemployment rates in Russia and Germany. Demidova and Signorelli (2012) investigate the determinants of youth unemployment in Russian regions, also taking into account the spatial effects. Guriev and Vakulenko (2012) and Kholodilin el al. (2012) study convergence among Russian regions in regional GDP per capita. Migration issues are investigated by Vakulenko (2013) along with the analysis of migration effect on wages, regional unemployment and per capita income. The unbalanced dynamics of Russian regions is also studied by Carluer and Sharipova (2004), who examined the convergence process of the Russian economy and confirm the regional divergence.

The Russian regions considered in our analysis are 75 (appendix A). According to the Russian Constitution, the Russian Federation consists of republics, krais, oblasts, cities of federal importance, an autonomous oblast and autonomous okrugs, all of which are equal subjects (regions) of the Russian Federation. By 2008 the number of federal subjects had decreased to 83 because of several mergers. Due to administrative reforms and lack of data, data on 8 regions are not included: Yamalo-Nenetsky Autonomous District - (62), Chukotka Autonomous District - (76), Dagestan, Republic of - (37), Ingushetia, Republic of - (38), Chechnya, Republic of - (36), Buryatia, Republic of - (65), Altai Territory - (63)[1]. The Map 1 represents the Russian regions.

[1] In 2014 Sevastopol and the Republic of Crimea became the 84th and 85th federal subjects of Russia.

Map 1. The Russian regions

Map 2. The Economic regions of Russia

These groups of federal subjects are also divided into twelve economic regions (see appendix A and Map 2) sharing the following characteristics: 1) common economic and social goals and participation in development programs; 2) relatively similar economic conditions and potential; 3) similar climatic, ecological, and geological conditions; 4) similar methods of technical inspection of new construction; 5) similar methods of conducting customs oversight; 6) overall similar living conditions of the population.

2) Indicators of the productive structure and labour market

We estimate a proxy for the labour market and productive structures of the regions by applying a dynamic multivariate factorial analysis. The method applied (STATIS: Structuration des Tableaux A Trois Indices de la Statistique, -in English: Structuring Three-way data sets in Statistics) (Escoufier, 1985; Escoufier, 1987) is well suited to the study of multidimensional phenomena like regional disparities because the regions (cases) can be analysed on the basis of a set of indicators (variables) that change over the years (time).

The STATIS methodology -see appendix for a more detailed description- consists in the analysis of a three-way data matrix $X_{I,JT}$ obtained from the temporal succession of data matrices ${}_tX_{i,j}$ of the same order, where i is the statistical unit and j the variable, both of them relative to the period t (i = 1, 2...I; j = 1, 2...J; t=1, 2...T).

The analysis moves through three phases: interstructure, compromise and infrastructure.

The purpose of the interstructure phase is to identify a suitable vectorial space smaller than *T,* where the *T* occasions (years) can be represented. To this end, examination is made of the matrix $I_{T,T}$, (also called the interstructure matrix), the column vectors of which are assumed as characteristic elements of each of the *T* occasions. In our case this is reduced to two dimensions but still maintains a good similarity to the initial representation.

In the compromise phase, a fictitious structure or synthesis matrix is identified which optimally summarizes the information contained in the *T* variance and covariance matrices. This structure, called 'compromise', is given by the matrix *W* obtained as a linear combination of the elements u_1 of the eigenvector of the matrix $I_{T,T}$ corresponding to the highest eigenvector and the matrices $\Gamma_T = \hat{X}\hat{X}$, where $\hat{X}$ is the deviation matrix.

The compromise phase consists in the estimation of a synthesis matrix which yields a representation, in the two-dimensional space identified, of the characteristic indicators and of the average positions of the regions in the time-span analysed.

The result of the intrastructure phase is a representation of the trajectories followed by the individual regions in the same period of time.

This dynamic multivariate method enables us to cluster regions year by year on the basis of a set of variables comprising labour market and income indicators, as well as indicators of the population structure and the structure of the productive sector. It is thus possible to study how the interaction between the labour market structure and economic growth changes over time, and also to analyse the dynamics of regions.

The variables used for this analysis are listed in Table 1 (the data are collected from the database provided by Federal State Statistics Service). They are indicators characteristic of the labour market and the production system (Wishlade and Yuill, 1997). Labour demand was measured by the employment rate (EMR), while the labour supply was measured by the labour-force participation rate (ACR). The Unemployment rate (UNR) was used as a proxy of the gap between labour demand and supply. The percentage of young population (YOU) was used as a measure of the demographic structure of the region.

The production system was represented by three variables corresponding to the percentages of employed persons in agriculture (AGR), industry (MAN), and traditional services (TRA). The other variables considered were Urban density (URB), as a proxy for the agglomeration factors of a region (Fujita M. et al., 2001; Krugman P.R., 1991), the per capita income (GDP), which is the indicator most frequently used to represent regional disparities, and the share of people with high education (SHE) as an indicator of Human Capital. Descriptive statistics of the variables are presented in the Appendix B.

Table 1. Variables used in the STATIS analysis

N	Proxy	Variable	Measure	Acr.
1	Regional Economic performance indicator	Per capita Gross Domestic Product	GDP per capita in price 2005 Correct for the Consumer' Purchasing Power	GDP
2	Agglomeratio n factors	Urban density	Share of Urban Population	URB
3	Labour Supply	Total activity rate	Active population/population aged over 15	ACR
4	Labour demand	Employment rate	employed/population aged over 15	EMR
5	Gap between Labour and Supply	Unemployment rate	Unemployed/Active population	UNR
6	Indicator of the demographic pressure	Share of population below 15 years	Population below 15 years/Population	YOU
7	Productive structure of the regional economy	Percentage employment in agriculture	Employed in agriculture/ total employed	AGR
8		Percentage employment in industry	Employed in industry/total employed	MAN
9		Percentage employment in traditional services	Employed in retail trade, hotels and non-market services /total employed	TRA
10	Human Capital indicator	Share people with high education	Population with tertiary education/population 15-64 aged	SHE

The time period considered for the analysis was seven years, from 2007 to 2013.

3) The analysis of the structure and the dynamics of the Russian regions

A global comparison between data tables is done using the RV coefficient (Escufier index) (see appendix D) representing an index of dissimilarity between years. The *RV* coefficients are non-negative and ranges between 0 and 1, and the closer *RV* is to 1 means the more similar the two data matrices *k* and *k´* are.

Trough the analysis of the coefficients we can conclude that the contiguous years are the closest ones. The most similar seem to be the 2010-2011 and the 2012.

Table 2: Matrix of the RV coefficients

	2007	2008	2009	2010	2011	2012	2013
2007	1.0000						
2008	.9581	1.0000					
2009	.9415	.9496	1.0000				
2010	.9203	.9308	.9498	1.0000			
2011	.9009	.9143	.9397	.9701	1.0000		
2012	.9054	.9209	.9369	.9574	.9689	1.0000	
2013	.8807	.8898	.9101	.9193	.9380	.9574	1.0000

Source: Our calculations on Russian data collected from official database (provided by Federal State Statistics Service)

In order to evaluate the goodness of the factorial representation yielded by construction of the compromise matrix, Table 3 shows the first three highest eigenvalues and the percentage of the total variance explained by the first three factorial axes.

Table 3. Eigenvalues and inertia percentages of the factorial axes

Axis	Eigenvalue	Variance explained	Cumulated variance explained
1	4.16198	40.62	40.62
2	2.12571	20.75	61.37
3	.941978	9.19	70.56

Source: Our calculations on Russian data collected from official database (provided by Federal State Statistics Service)

To be noted first is that 40,6% of the variance is explained by the first factor, and 20,8% by the second, for a total of 61,4% of the variance expressed by the set of all the variables. In other words, the first factor alone explains more than a third of the total variability, while the first three factors jointly explain almost 70,6%. Consequently, the reduction of the phenomenon’s variability, obtained by representing it in a two or three-dimensional space, is a meaningful synthesis of the information considered.

Figures 1. and 2. show, respectively on the factorial plane generated by the first two and by the first and third principal components, the positions of the average annual value of each of the ten characteristic indicators considered.

Fig. 1. Position of the characteristic indicators on the first and second factorial plan

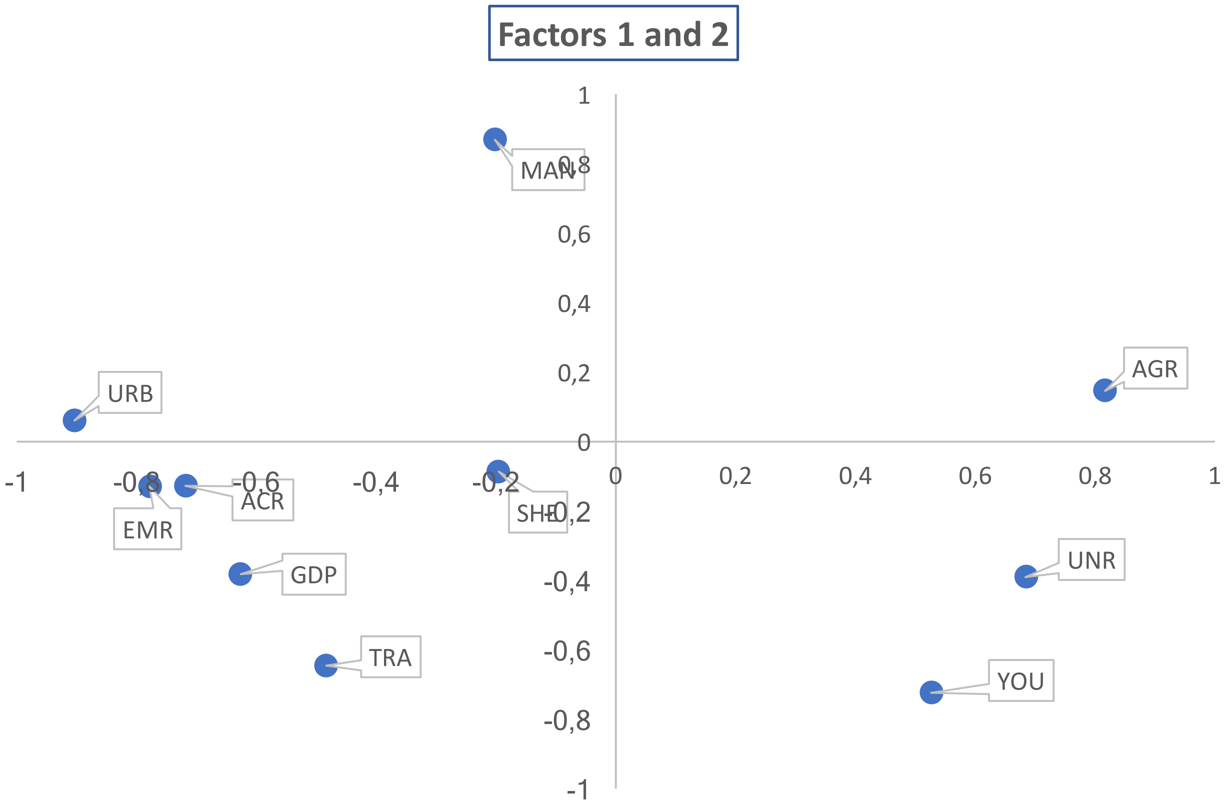
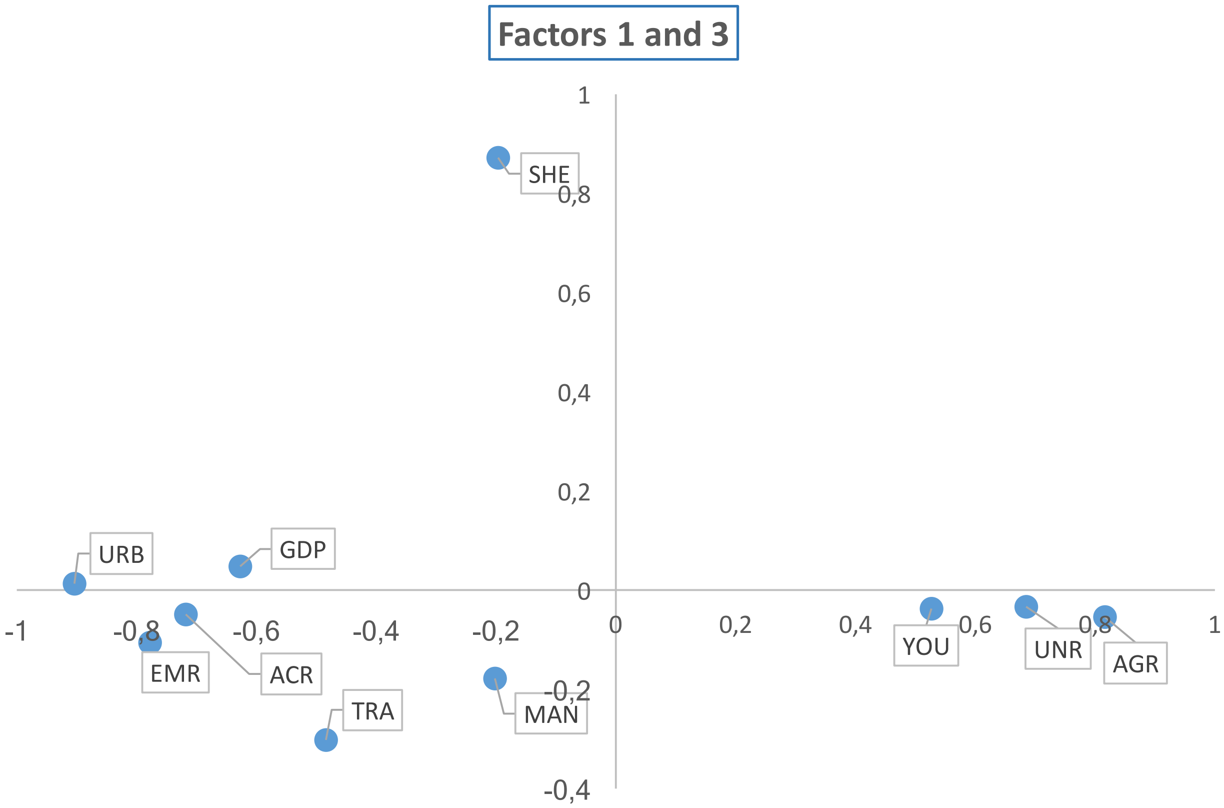


Fig. 2. Position of the characteristic indicators on the first and third factorial plan



In order to interpret the factors, we may refer to Table 4, which shows that minimum and maximum values of the correlations between the variables and the factorial axes. It can be seen that the variables most closely correlated with the first factor are the urbanization rate (URB), the employment rate (EMR), the activity rate (ACR), the per capita income (GDP), and the percentage of employment in traditional services (TRA) on the one hand (negative quadrants) and the unemployment rate (UNR), the percentage of young people (YOU) and the share of employment in the agricultural sector (AGR) on the other (positive quadrants). In other words, along the first axis one observes a clear polarization between a rich labour market structure and indicators of high urbanization and those relative to high unemployment, demographic structure, high share of young people, and the presence of agricultural employment.

Along the second axis one observes a close correlation among, on the one hand (positive quadrant), the high percentage of employment in industry (MAN), opposed to a mix of variables which are the percentage of young people (YOU) and the unemployment rate (UNR) positioned in the fourth quadrant and the percentage of employment in traditional services (TRA) and per capita income (GDP) positioned in the third quadrant. In this case, the second axis identifies in a marked manner the phenomena representing variables located in the positive quadrant, namely the industrial structure, and variables linked to the population structure of the less developed areas (quadrant IV) and to the service and to high income variables in the richest areas (quadrant III).

The third axis is identified in a manner marked only in the positive quadrant by the indicator of human capital (SHE).

Table 4. Correlations between the variables and the factorial axes (minimum and maximum period values)

Factor 1			Factor 2			Factor 3		
	Max	Min		Max	Min		Max	Min
URB	-0.91	-0.90	YOU	-0.73	-0.72	TRA	-0.32	-0.27
EMR	-0.81	-0.76	TRA	-0.70	-0.56	MAN	-0.19	-0.16
ACR	-0.74	-0.69	UNR	-0.50	-0.33	UNR	-0.15	-0.01
GDP	-0.65	-0.58	GDP	-0.40	-0.3	EMR	-0.15	-0.06
TRA	-0.51	-0.43	SHE	-0.18	0.00	ACR	-0.14	0.00
SHE	-0.29	-0.13	ACR	-0.17	0.09	AGR	-0.09	-0.02
MAN	-0.22	-0.17	EMR	-0.17	-0.09	YOU	-0.06	0.02
YOU	0.55	0.50	URB	0.06	0.05	URB	0.03	0.01
UNR	0.74	0.66	AGR	0.13	0.17	GDP	0.09	0.13
AGR	0.83	0.80	MAN	0.88	0.86	SHE	0.91	0.81

Source: Our calculations on Russian data collected from official database (provided by Federal State Statistics Service)

In conclusion, the three factorial axes represent certain characteristics of the labour market and the productive structure. The first factor (**FF**) can be interpreted as being a proxy for the ‘bad’ performance of the labour market. It should be pointed out that the variable has an opposite sign with respect to the development indicator: the regions that achieve a good performance in terms of activity rate and employment rate, higher per capita income levels and urbanization have negative values for this factor. By contrast, those regions that have high unemployment rates, high percentages of employed in agriculture and high percentage of youth population, have positive values. The second factor (**SF**) is mainly explained by the industrialisation index in the positive quadrant (first quadrant) and by the transport and gdp indices (third quadrant) and by youth unemployment and unemployment rate in the fourth quadrant. The third factor (**TF**) is mainly explained by the Human Capital proxy in the positive quadrant.

Figures 1A and 2A in the appendix C show the Russian regions placed, respectively, along the first two factorial axes and the first and the third axe. In order to better read the figures, we have limited (figure 3 and 4) the regions that overlap the position of the variables explaining the factorial axes by boxes of different colours. For example, Figure 3 shows the positions of the regions along the first factor and the second factor and in particular into the blue box there are regions associated with the variables characterizing the left part of the first factor that are the dynamic labour markets producing high levels of employment and participation, urbanization, high GDP and presence of transport infrastructure such as Moscow Region, Moscow, Murmansk, Saint-Petersburg, Samara, Tumen, Magadan, Sakhalin. The red box contains regions associated to the variables characterizing the right side of the first factor i.e. the high unemployment rate, a high percentage of employment in agriculture and a strong demographic pressure as: Tambow, Rep. of Adygea, Republic of Kabardino, Republic of Kalmykia, Republic of Karachevo, Republic of Nothern Osetia, Krasnodar, Stavropol, Astrakhan, Rostov, Republic of Bashkortosan, Kurgan, Republic of Altay, Republic of Buryatia, Republic of Tyva, Altay Territory, Jewish Autonomos. On the other hand, the green circle includes regions associated with the variable characterizing the upper side of the second factor i.e. the high presence of employed in industry such as Vladimir, Ivanovo, Kaluga, Tula, Ulyanovk. The azure box contains regions associated with a mix of variables located in the lower side of the second factor (Rep. of Karelia, Rep. of Komi, Murmansk, Rep. of Kabardino, Rep. of Karacjaevo, Tumen, Rep. of Altay, Rep. of Buryatia, Rep. of Tyva, Rep. of Sakha, Amur, Magadan, Sakhalin). Finally, the figure 4 synthetizes the position of the main regions along the third factor. In particular, on the upper side we have regions associated with the main variable explicative of the third factor that is the index of human capital that are Moscow, Saint-Petersburg, Volgograd. On the other side, the red box and the green box contain regions associated with mix of variables of difficult interpretation but also with a very low correlation.

Fig. 3. Position of the regions on the first and second factorial plan

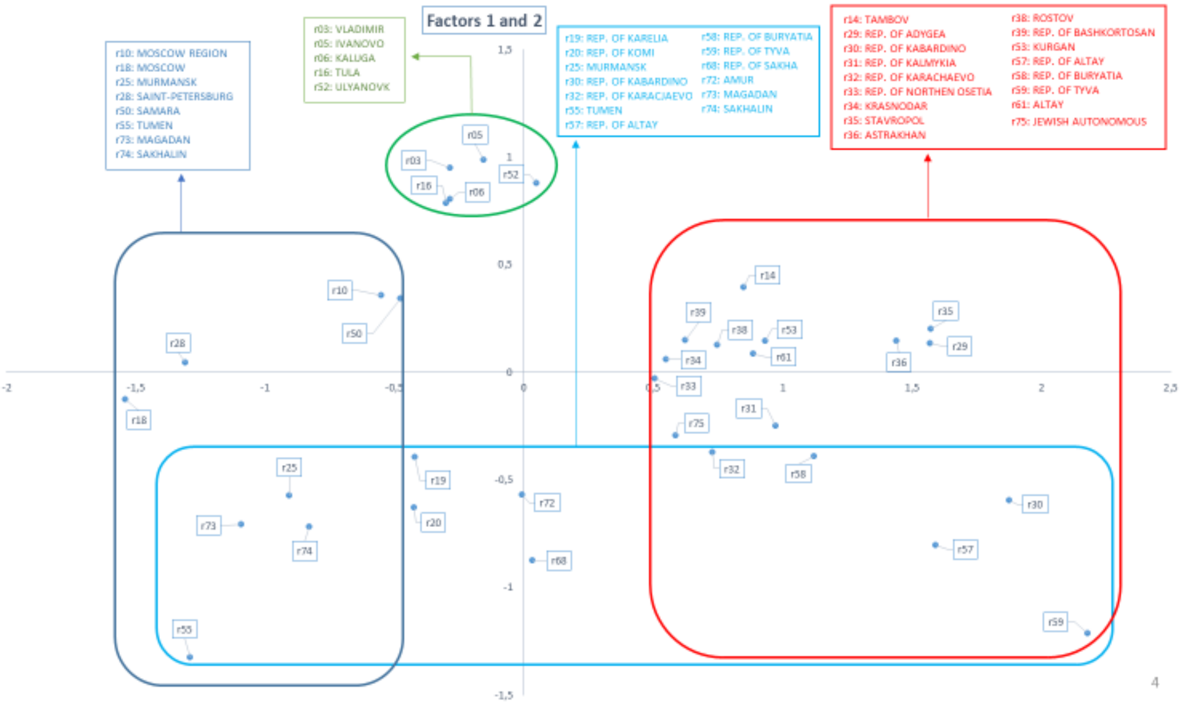
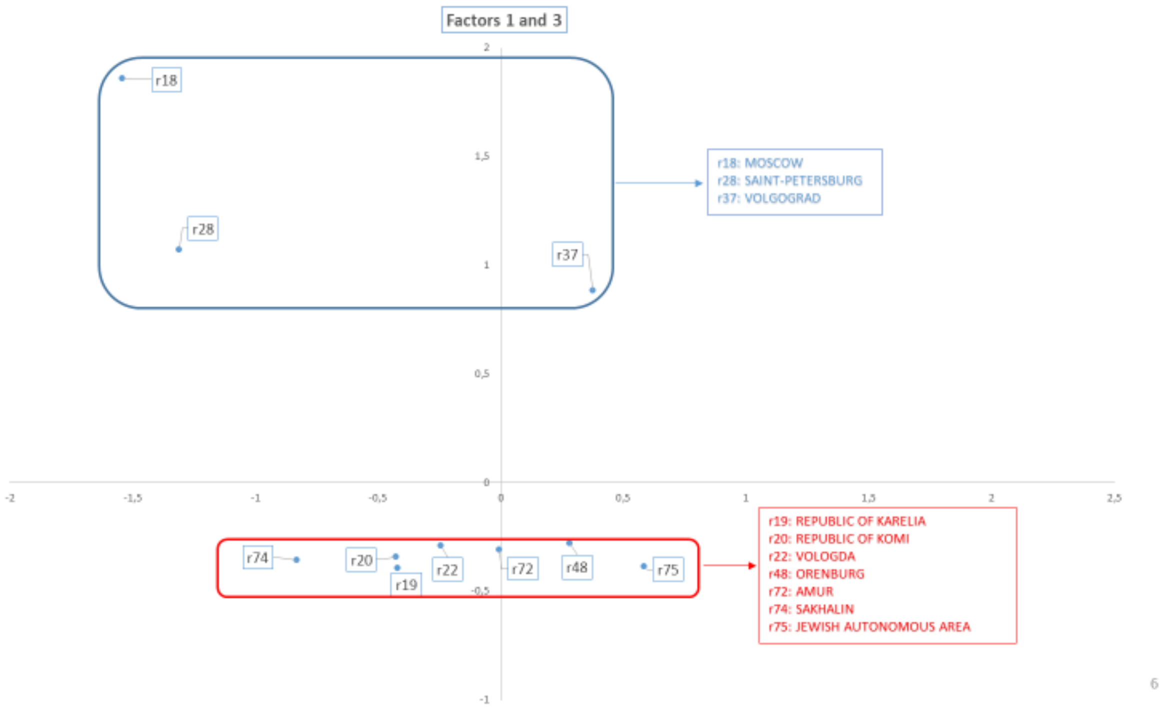
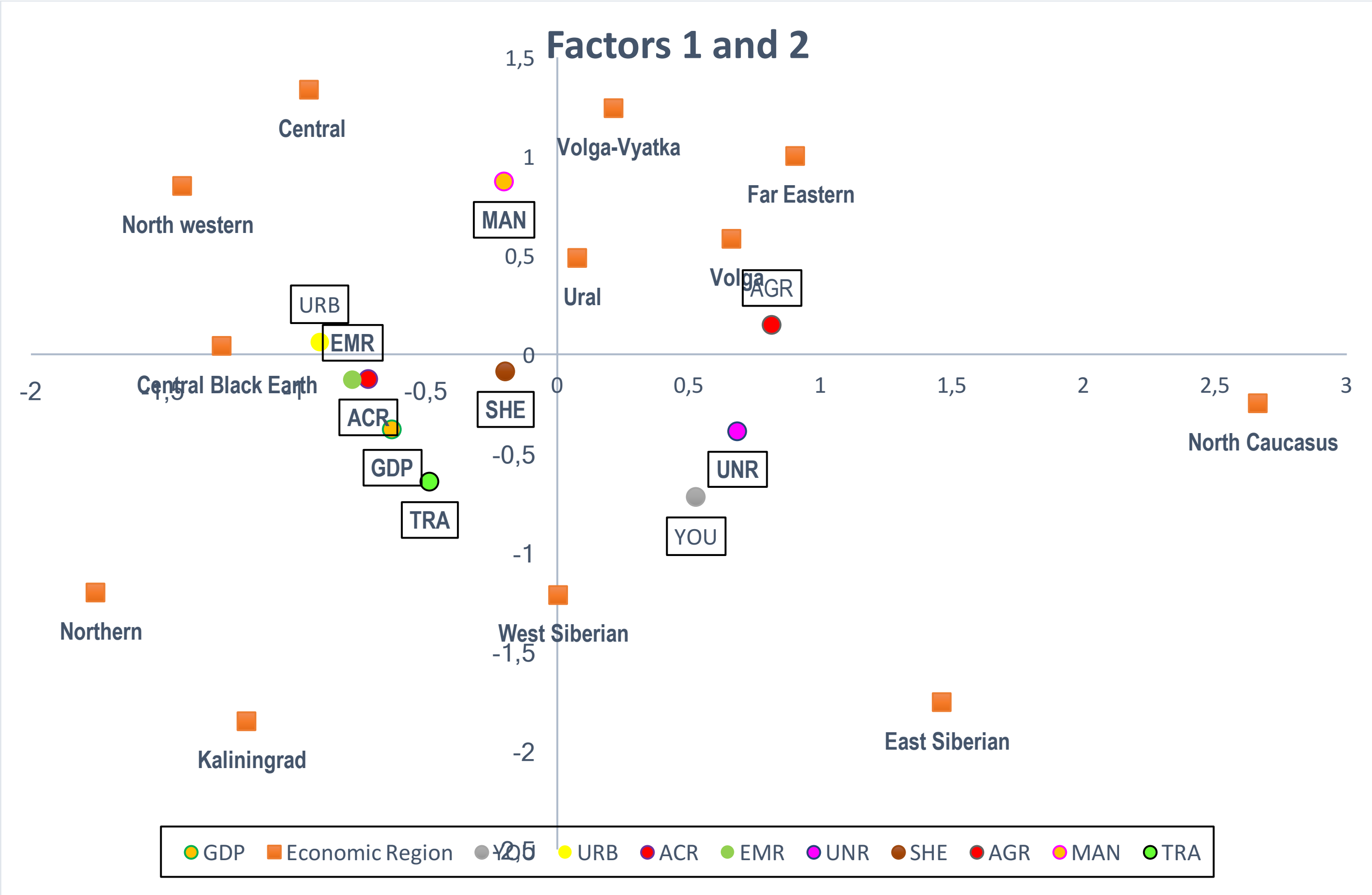


Fig. 4. Position of the regions on the first and third factorial plan



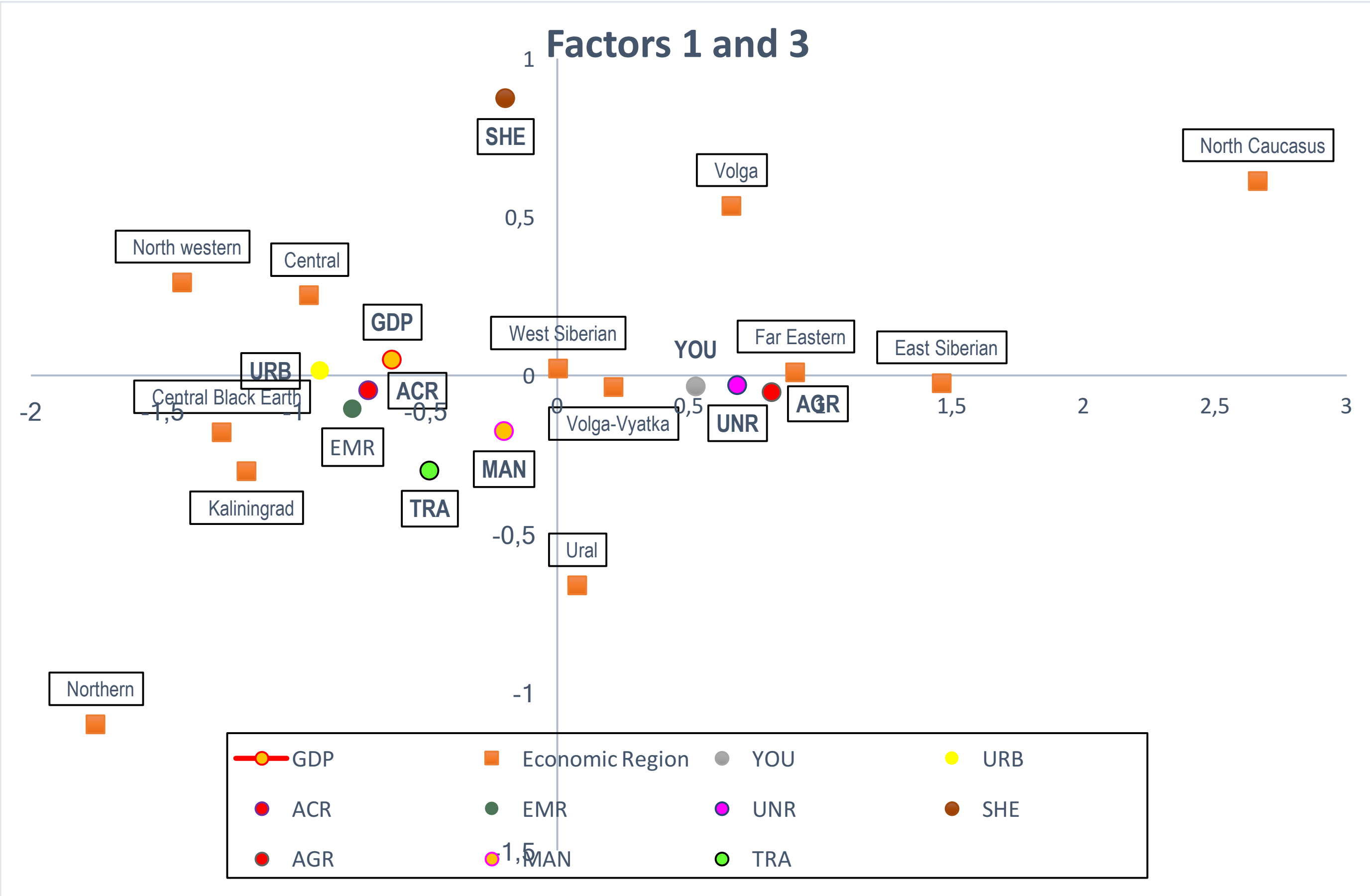
More interestingly, the following two figures show the position along the two factorial axes of the twelve macro regions. As you can see, the macro-regions of central and northern Europe are located in the first and third quadrant, that is they are characterized by a high economic development and specialised or by a high percentage of employment in the manufacturing sector (Central, North Western and Central Black Heart, Volga-Vyatka) or by a high presence of the service sector (Northern, Kalingrad). Volga (as well as Far Eastern federal district) is characterised by high share of employment in agriculture and industry (in particular mining), whereas Ural's main characteristic is the high share of employment in industry. The North Caucasus region and the Siberian regions are, instead characterised by a high unemployment and by a demographic pressure.

Fig. 5. Position of the macro regions on the first and second factorial plan



As regards the position of the macro regions along the third factor, it is possible to note that the regions placed in the quadrants with positive values of the third factor, i.e. that are associated with the proxy variable of human capital (high share of population with tertiary education) (SHE), are North Western and Central regions (more developed) and the less developed regions of Volga and of North Caucasus.

Fig. 6. Position of the macro regions on the first and third factorial plan



3.1 Intrastructure-Trajecories

A further result of the intrastructure analysis concerns the temporal trajectories followed by individual regions along the factorial axes and highlighting certain characteristics of the regional dynamic. A summary of these phenomena is provided by Tables 5-8, which show – for each year and only for the first two factors – the sums of the square of the distances between the individual regions and the factorial axis, weighted for the region's contribution to formation of that axis. In this way greater importance is given to the paths followed by the regions making the greatest contribution to defining the factor. The distances have been separately calculated for all regions and for the twelve macro-regions.

Table 5. Weighted average annual distance of the regions from the first factorial axis.

Year	Central	Northern	Volga	North Caucasus	North western	Ural	East Siberian	Far Eastern	Central Black Earth	West Siberian	Volga-Vyatka	Kaliningrad	TOT
2007	0,4375	0,2905	0,6396	3,7987	0,0642	0,1164	5,9877	0,8143	0,0632	2,9013	0,0281	0,0309	15,1724
2008	0,6097	0,2113	0,5773	3,7818	0,0816	0,1470	6,2209	0,9216	0,0894	2,5400	0,0317	0,0319	15,2443
2009	0,5521	0,2780	0,6690	4,1380	0,0648	0,1175	7,3346	0,8536	0,0653	2,7709	0,0270	0,0212	16,8920
2010	0,4599	0,3208	0,3888	3,6433	0,0809	0,1302	7,4861	1,0302	0,0584	2,7425	0,0317	0,0135	16,3862
2011	0,4496	0,2823	0,3906	3,4067	0,0765	0,1060	8,2483	1,0633	0,0444	2,4977	0,0231	0,0045	16,5929
2012	0,5243	0,1722	0,4222	3,7026	0,0893	0,1115	5,9345	1,1181	0,0415	2,6754	0,0242	0,0143	14,8301
2013	0,8516	0,2151	0,4835	4,1537	0,0432	0,1109	9,2087	1,1424	0,0391	2,9632	0,0369	0,0140	19,2622

Table 6. Index numbers

	Central	Northern	Volga	North Caucasus	North western	Ural	East Siberian	Far Eastern	Central Black Earth	West Siberian	Volga-Vyatka	Kaliningrad	TOT
2007	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
2008	139,36	72,74	90,26	99,56	127,15	126,31	103,89	113,18	141,47	87,55	112,88	103,43	100,47
2009	126,20	95,70	104,60	108,93	100,91	100,93	122,49	104,83	103,27	95,51	96,21	68,69	111,33
2010	105,11	110,41	60,79	95,91	125,91	111,83	125,02	126,52	92,43	94,53	112,89	43,69	108,00
2011	102,77	97,16	61,07	89,68	119,12	91,07	137,75	130,57	70,29	86,09	82,01	14,62	109,36
2012	119,84	59,28	66,01	97,47	139,00	95,81	99,11	137,31	65,63	92,22	86,24	46,33	97,74
2013	194,66	74,05	75,59	109,34	67,23	95,25	153,79	140,30	61,81	102,14	131,40	45,36	126,96

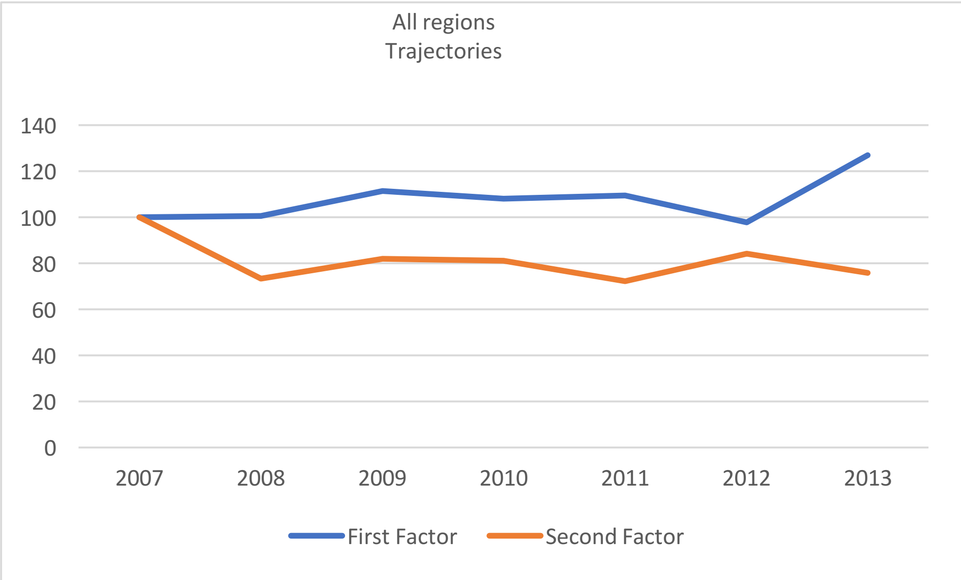
Table 7. Weighted average annual distance of the regions from the second factorial axis

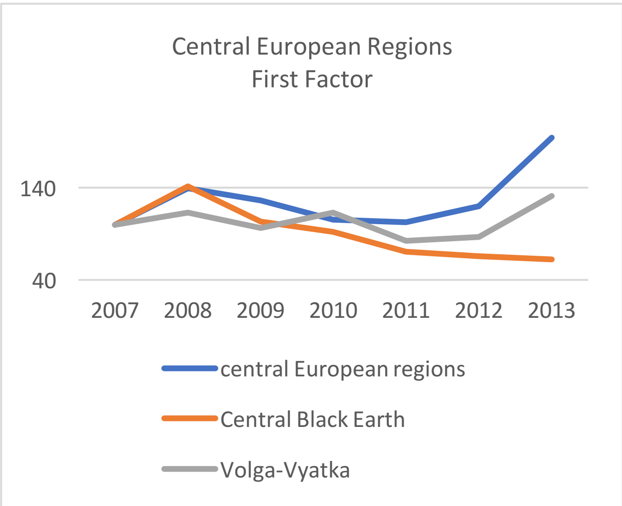
	Central	Northern	Volga	North Caucasus	North western	Ural	East Siberian	Far Eastern	Central Black Earth	West Siberian	Volga-Vyatka	Kaliningrad	TOT
2007	2,3750	0,0739	0,1358	0,0146	1,0169	0,4149	0,2743	105,2643	0,0704	5,3734	0,4248	0,0002	115,4384
2008	0,8512	0,0163	0,0414	0,0162	0,5128	0,9486	0,0137	77,0424	0,0328	4,2427	0,4348	0,0007	84,1538
2009	0,8596	0,0249	0,0403	0,0263	0,5476	1,2572	0,0135	86,1607	0,0393	4,5050	0,6819	0,0004	94,1567
2010	0,8397	0,0120	0,0343	0,0288	0,4218	1,3095	0,0229	85,3756	0,0305	4,6208	0,4734	0,0005	93,1699
2011	0,6581	0,0163	0,0344	0,0769	0,4018	1,2948	0,0195	76,0746	0,0247	4,1077	0,3706	0,0004	83,0801
2012	0,7134	0,0157	0,0382	0,1026	0,5020	1,7632	0,0128	88,5169	0,0381	4,7415	0,4549	0,0002	96,8997
2013	0,2395	0,0155	0,0264	0,0252	0,4378	2,2322	0,0129	79,8967	0,0396	4,2140	0,3406	0,0002	87,4805

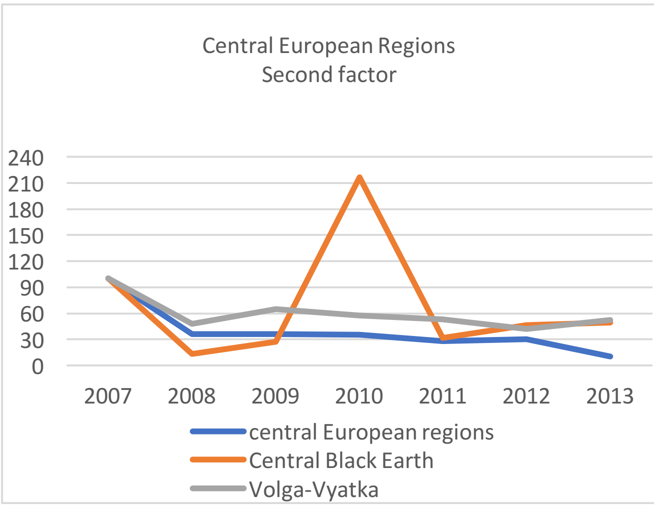
Table 8. Index numbers

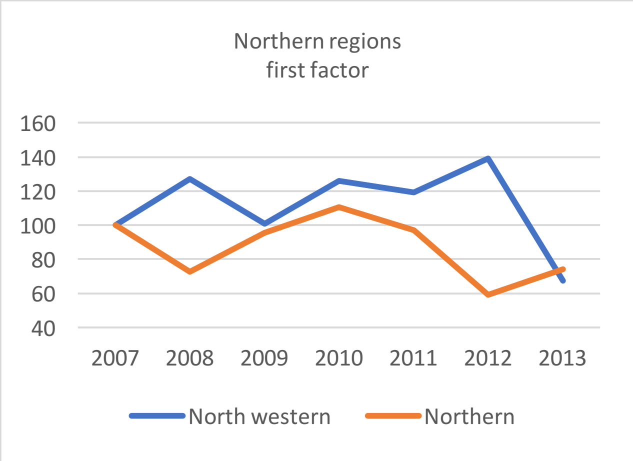
Year	Central	Northern	Volga	North Caucasus	North western	Ural	East Siberian	Far Eastern	Central Black Earth	West Siberian	Volga-Vyatka	Kaliningrad	TOT
2007	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
2008	35,84	22,06	30,50	111,34	50,43	228,64	5,00	73,19	46,63	78,96	102,35	323,49	73,19
2009	36,19	33,66	29,70	180,76	53,86	303,01	4,93	81,85	55,76	83,84	160,50	185,24	81,85
2010	35,36	16,30	25,27	197,45	41,49	315,63	8,37	81,11	43,30	85,99	111,43	213,79	81,11
2011	27,71	22,11	25,32	528,17	39,52	312,08	7,12	72,27	35,16	76,45	87,24	194,63	72,27
2012	30,04	21,26	28,11	704,88	49,37	424,98	4,68	84,09	54,16	88,24	107,08	103,79	84,09
2013	10,09	20,96	19,41	172,94	43,05	538,01	4,72	75,90	56,32	78,42	80,16	69,38	75,90

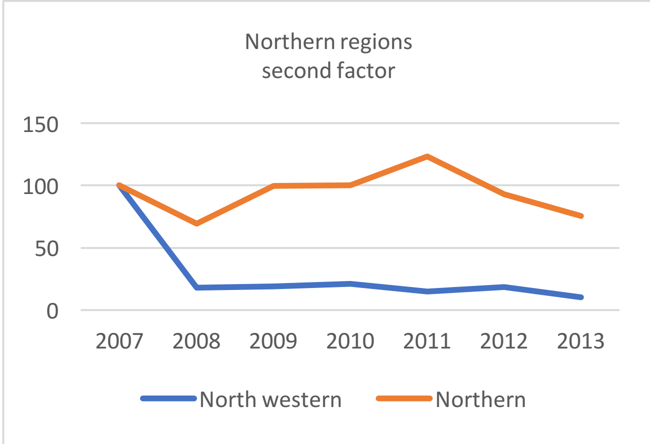
The following Figures synthetize the dynamic of the number indices. A first general phenomenon to be observed is that whilst for factor 1 the total distance slightly increased during the period for all the groups of regions considered, it diminished for factor 2. This seems to indicate that the regions gradually moved closer to the phenomena characterizing the second factor. A second feature to be noted is that Macro Regions that seem to go against the trend with respect to the first axis, i.e. are reducing the distance, are The Northern region the west Siberian region and the Volga region. On the other hand, the regions that seem to go against the trend with respect to the second axis, i.e. are augmenting the distance, are the North Caucasus Regions and the East Siberian Region. The third feature to stress is that the pattern of the distances does not seem cyclical.

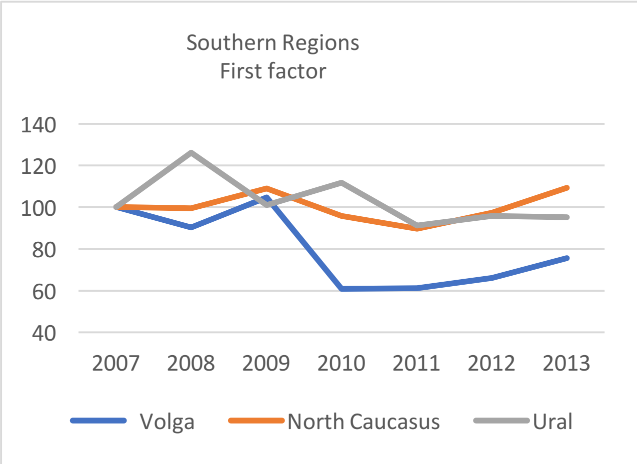


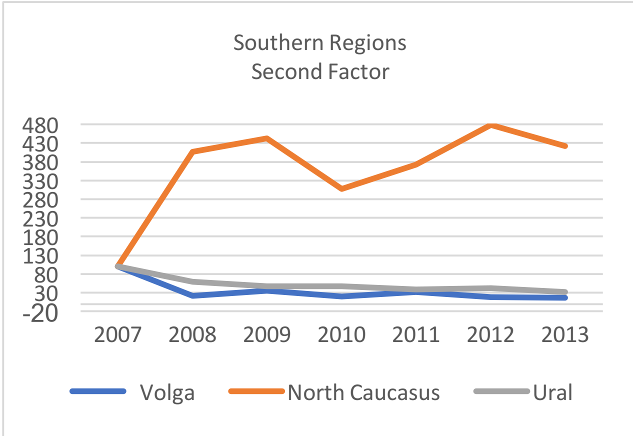












Summary and conclusions

The results of the analysis confirm the thesis of those who contend the Russian regions have a diversified reality influenced by structural phenomena concerning labour market characteristics, sectoral composition, and localization factors. This makes unlikely that integration processes – although accelerated by the enlargement of markets and their greater efficiency – will give rise to the hope for levelling of economic development in the near future. The main reason for regional differences still seems to be the composition and structure of labour market and industry. To be noted in particular is the marked contrast between the Central and Northern European regions characterized by more flexible labour markets and high employment rates and the Siberian and Southern East regions characterised by high rates of structural unemployment. However, there are other phenomena responsible for regional disparities in Russia: localization factors (large conurbations, transport hubs, and *tourism*) which foster the development of connected service activities, and the presence of a solid industrial base accompanied by high levels of income and employment. These factors are associated with regions which are more territorially dispersed and therefore unlikely to form regional clusters, whilst, by contrast, industrialization phenomena are distributed across a transnational area formed by contiguous regions. The dynamic analysis has shown not so much convergence as slow change in the structural characteristics that differentiate the regions, where localization factors and sectoral composition will probably be more influential in the future. Moreover, the peripheral regions seem to be more markedly characterized by structural differences than are the core regions.

References

Amendola, F.E.Caroleo and G. Coppola, (2006) "Regional Disparities in Europe" in F. E. Caroleo and S. Destefanis (eds) *The European Labour Market: Regional Dimensions*, Physica Verlag: Heidelberg, 9-31.

Boldrin, L. and F. Canova, (2001) "Inequality and Convergence in Europe's Regions. Reconsidering European Regional Policies", *Economic Policy*, 16, 207-253.

Barro R.J. and X. Sala-i-Martin, (1991), "Convergence across States and Regions", *Brookings Papers on Economic Activity* 1, 107-182.

Blanchard O. J. and L. F.Katz, (1992), "Regional Evolutions", *Brooking Papers on Economic Activity*, 1, 1-61.

Caroleo F. E. and G. Coppola, (2006) "The Impact of the Institutions on Regional Unemployment Disparities in Europe," D.E.S. (Department of Economic Studies), University of Naples "Parthenope", *Discussion Papers* n. 4.

Caroleo F.E. and F. Pastore (2010) "Introduction" in: Caroleo F.E and F. Pastore (eds) *The Labour Market Impact of the EU Enlargement: A new Regional Geography of Europe?*.Physica-Verlag: Heidelberg, 1-13.

F. E. Caroleo and F. Pastore, (2009) "Structural Change and Regional Imbalances in Transition" in E. Marelli, M. Signorelli (eds) (2009), *Growth and Structural Features of Transition*, Palgrave Macmillan: London, 260-276.

Carluer, F., & Sharipova, E. (2004). The unbalanced dynamics of Russian regions: towards a real divergence process. East-West Journal of Economics and Business, 7(1), 11-37.

D'Ambra L., (1985) "Alcune estensioni dell'analisi in componenti principali per lo studio di sistemi evolutivi, Uno studio sul commercio internazionale dell'elettronica", *Ricerche Economiche n.2*, 233-260.

Decressing, J. and A. Fatàs, (1995) "Regional Labor Market Dynamics in Europe", *European Economic Review. n. 3*, 1627-1655,

Elhorst J.P., (2003) "The Mystery of Regional Unemployment Differentials: a Survey of Theoretical and Empirical Explanations"*, Journal of Economic Surveys,* Vol.17(5), 709-748.

Escoufier Y., (1985) "Statistique et analyse des données", *Bulletin des Statisticiens Universitaires, 10*.

Escoufier Y., (1987) "Three-mode data analysis: the Statis method*",* in: Fichet, B., Lauro, C. (eds.) *Methods for Multidimensional Data Analysis*, pp.259–272. ECAS: Napoli.

Fachin S. and M. Vichi, (1994) “Deindustrializzazione, specializzazione o ristrutturazione? Una analisi multiway in matrici fattoriali dell’evoluzione dell’industria manifatturiera italiana dal 1971 al 1983”, *Politica Economica,* 10, 373-404.
Ickes B. W. and Ofer G. (2006), “The Political Economy of Structural Change in Russia”, *European Journal of Political Economy*, Vol. 22(2), 409-434.

Ivanova V.I. (2015) “Convergence Of Prices On The Grain Market: The Historical Aspect”, Prostranstvennaya Ekonomika = Spatial Economics, no 3, pp. 34–56. (In Russian).

Kolomak E. A. (2010) “Spatial Externalities as a Source for Economic Development”, Region: economics and sociology, 2010, no 4, pp.73-87. (In Russian).

Lugovoy O., Dashkeev V., Mazaev I. (2007) “Analysis of Economic Growth in Regions: Geographical and Institutional Aspect”. Consortium on Applied Economic Research. The Canadian International Development Agency. Moscow: Institute for the Economy in Transition, 164 p. (In Russian).

Marelli E. and M. Signorelli, (2010) “Economic growth and structural features of transition: Theoretical framework and general overview” in: Marelli E. and M. Signorelli (eds) *Economic Growth and Structural Features of Transition*. Palgrave Macmillan: Basingstoke, 11-38.

Marelli E., Patuelli R. and M. Signorelli, (2012) “Regional Unemployment in the EU before and after the Global Crisis”, *Post-Communist Economies*, Vol. 24(2), 155-175

Niebuhr A. (2002), “Spatial Dependence of Regional Unemployment in the European Union”, *HWWA Discussion Paper*, n. 186.

Obstfeld M. and G. Peri (1998), “Regional non-adjustment and fiscal policy”, *Economic Policy*, 205-47.

Pastore F., (2912) *“Primum vivere* . . . industrial change, job destruction and the geographical distribution of unemployment”, *IZA Journal of European Labor Studies* 1:7.

Rivadeneira F. J., Figueiredo A. M. S., Figueiredo F. O. S., Carvajal S. M. and R. A. Rivadeneira, (2016) “Analysis of well-being in oecd countries through statis methodology”, *Holos*, 32, (7), 335-351.

Semerikova E.V., Demidova O.A. (2015) “Analysis of Regional Unemployment in Russia and Germany: Spatial-Econometric Approach”. Prostranstvennaya Ekonomika = Spatial Economics, no. 2, pp. 64-85. (In Russian).

Tassinari G. and M. Vichi, (1994) “La dinamica economica dei paesi avanzati negli anni ottanta: Riflessioni sulle traiettorie risultanti dalle analisi delle matrici a tre vie”, *Giornale degli Economisti e Annali di Economia,* 1-3.

Vakulenko E.S. (2013) “Does Migration Lead to Regional Convergence in Russia?”, Vestnik NSUEM, T. 4, pp. 239–264. (In Russian).

Wishlade F. and Y. Douglas, (1997), “Measuring disparities for area designation purposes: issues for the European Union”, *Regional and Industrial Policy Research Paper,* n. 24, European Policies Research Centre University of Strathclyde.

Appendix A. Regions

According to the Russian Constitution, the Russian Federation consists of republics, krais, oblasts, cities of federal importance, an autonomous oblast and autonomous okrugs, all of which are equal subjects of the Russian Federation. By 2008 the number of federal subjects had decreased to 83 because of several mergers. In 2014 Sevastopol and the Republic of Crimea became the 84th and 85th federal subjects of Russia.

These groups of federal subjects are also divided into twelve economic regions —sharing the following characteristics:

- Common economic and social goals and participation in development programs;
- Relatively similar economic conditions and potential;
- Similar climatic, ecological, and geological conditions;
- Similar methods of technical inspection of new construction;
- Similar methods of conducting customs oversight;
- Overall similar living conditions of the population.

Central Black Earth		
	Belgorod region	R01
	Kursk region	R08
	Lipetsk region	R09
	Tambov region	R14
	Voronezh region	R04
Central		
	Brjansk region	R02
	Federal City of Moscow	R18
	Ivanovo region	R05
	Jaroslavl region	R17
	Kaluga region	R06
	Kostroma region	R07
	Moscow region	R10
	Orel region	R11
	Rjazan region	R12
	Smolensk region	R13
	Tula region	R16
	Tver region	R15
	Vladimir region	R03
East Siberian		
	Irkutsk region	R63
	Krasnoyarsk Territory	R62
	Republic of Buryatia	R58
	Republic of Khakassia	R60
	Republic of Tyva	R59
	Zabaykalsky Krai	nd
Far Eastern		
	Amur region	R72
	Chukotka Autonomous Okrug	nd
	Jewish autonomous area	R75
	Kamchatka territory	R69
	Khabarovsk Territory	R71
	Magadan region	R73
	Primorsky Territory	R70
	Republic of Sakha (Yakutia)	R68
	Sakhalin region	R74
Kaliningrad		
	Kaliningrad region	R23
North Caucasus		
	Krasnodar Territory	R34

	Republic of Adygea	R29
	Republic of Crimea	nd
	Republic of Dagestan	nd
	Republic of Ingushetia	nd
	Republic of Kabardino-Balkaria	R30
	Republic of Karachaevo Cherkessia	R32
	Republic of Northen Osetia – Alania	R33
	Republic of Sevastopol	nd
	Rostov region	R38
	Stavropol Territory	R35
Northern		
	Arkhangelsk region	R21
	Murmansk region	R25
	Nenets Autonomous Okrug	nd
	Republic of Karelia	R19
	Republic of Komi	R20
	Vologda region	R22
Northwestern		
	Federal city of St. Petersburg	R28
	Leningrad region	R24
	Novgorod region	R26
	Pskov region	R27
Ural		
	Chelyabinsk region	R56
	Kurgan region	R53
	Orenburg region	R48
	Perm territory	R45
	Republic of Bashkortostan	R39
	Republic of Udmurtia	R43
	Sverdlovsk region	R54
Volga		
	Astrakhan region	R36
	Penza region	R49
	Republic of Kalmykia	R31
	Republic of Tatarstan	R42
	Samara region	R50
	Saratov region	R51
	Ulyanovsk region	R52
	Volgograd region	R37
Volga-Vyatka		
	Kirov region	R46
	Nizhny Novgorod region	R47

	Republic of Chuvashia	R44
	Republic of Marii El	R40
	Republic of Mordovia	R41
West Siberian		
	Altay Territory	R61
	Kemerovo region	R64
	Khanty–Mansi Autonomous Okrug	nd
	Novosibirsk region	R65
	Omsk region	R66
	Republic of Altay	R57
	Tomsk region	R67
	Tumen region	R55
	Yamalo-Nenets Autonomous Okrug	nd

Data on the following 8 regions are not included in our analysis due to

1) administrative reform

2) some regions (for example, Chechnya) suffer from lack of data

Yamalo-Nenetsky Autonomous District - (62)

Chukotka Autonomous District - (76)

Dagestan, Republic of - (37)

Ingushetia, Republic of - (38)

Chechnya, Republic of - (36)

Buryatia, Republic of - (65)

Altai Territory - (63)

Chechnya

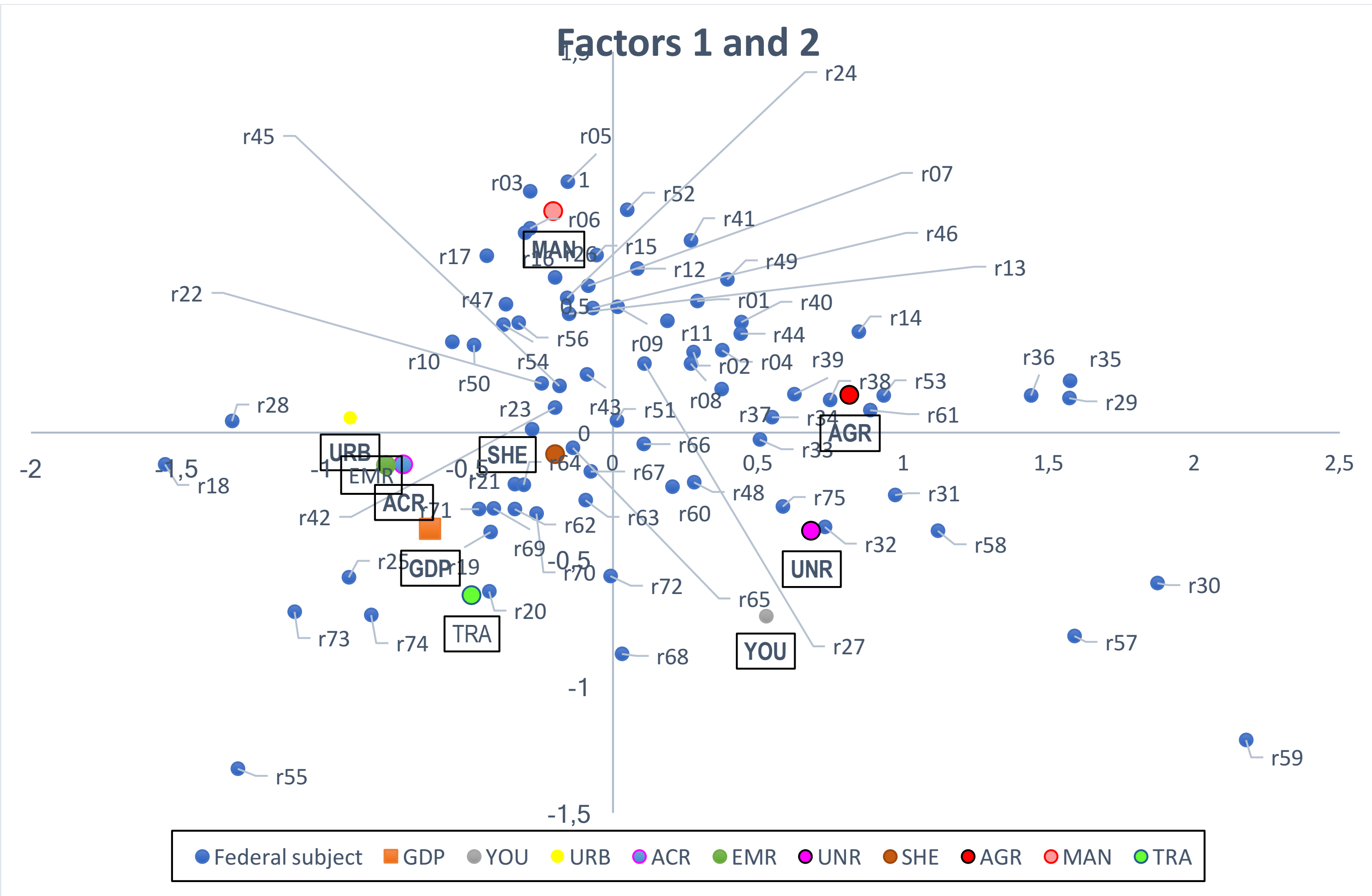
Appendix B Descriptive statistics

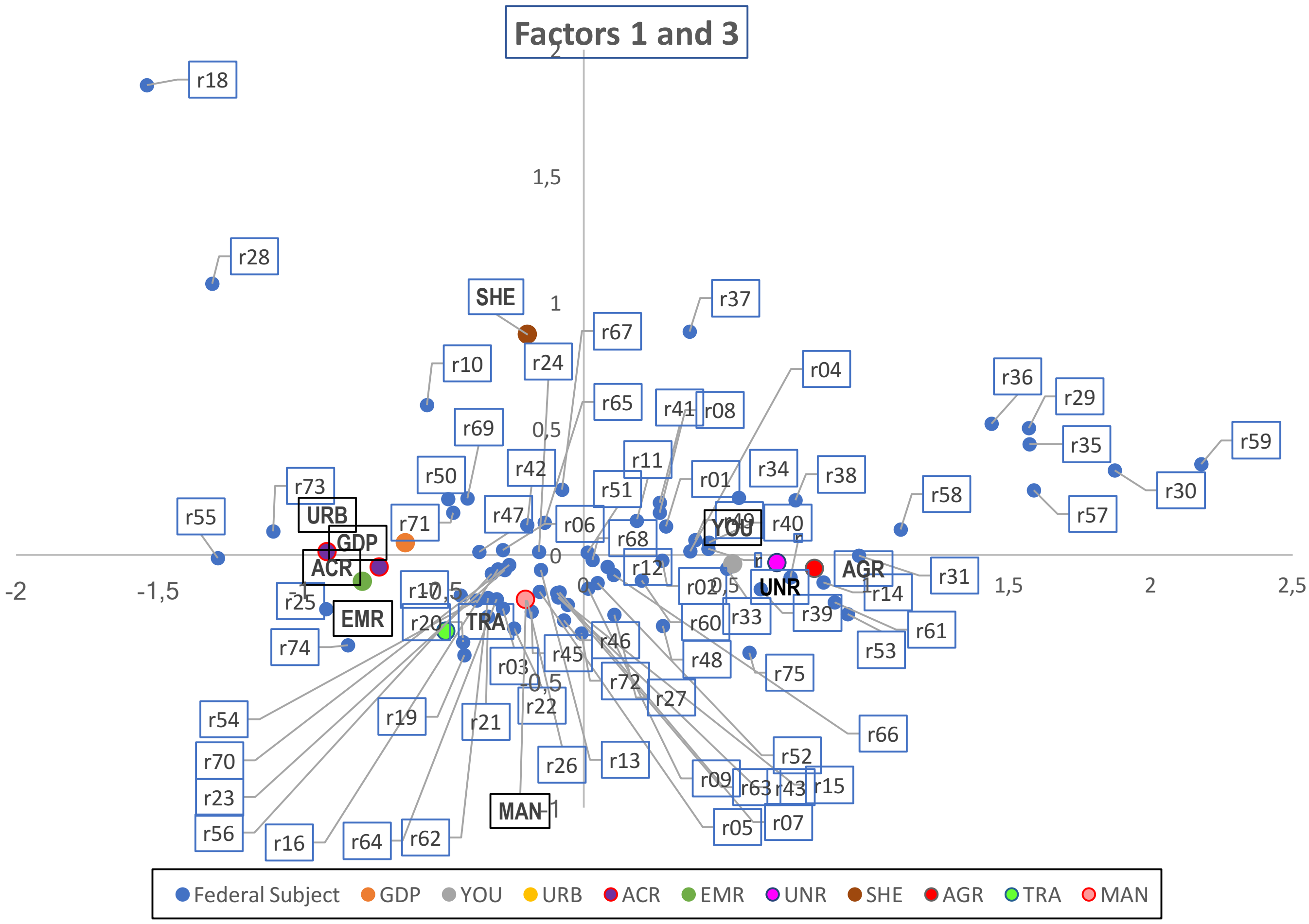
2005						
VARIABLES	N	mean	Var	sd	min	max
YOU	75	16.79	16.79	2.698	12.70	29.50
URB	75	69.82	69.82	12.22	26.20	100
GDP	75	974.9	974.9	643.2	379.5	5,473
ACR	75	65.74	65.74	3.053	57.80	73.50
EMR	75	46.58	46.58	4.482	34.42	57.40
UNR	75	8.193	8.193	3.563	0.800	23.90
SHE	75	22.12	22.12	5.088	14.30	43.80
AGR	75	13.27	13.27	5.531	0.200	26.70
MAN	75	16.73	16.73	6.117	4.100	30.70
TRA	75	15.30	15.30	2.976	9	23.80
2006						
YOU	75	16.40	16.40	2.650	12.40	28.90

URB	75	69.92	69.92	12.29	26.70	100
GDP	75	1,104	1,104	680.7	457.6	5,714
ACR	75	66.11	66.11	3.103	58.60	72.80
EMR	75	47.13	47.13	4.651	34.59	56.75
UNR	75	7.687	7.687	3.730	1.600	20.80
SHE	75	22.93	22.93	4.914	16.70	48.50
AGR	75	12.86	12.86	5.456	0.200	26.10
MAN	75	16.43	16.43	6.116	2.900	30.90
TRA	75	15.55	15.55	2.964	9.300	23.60
2007						
YOU	75	16.25	16.25	2.678	12.20	29
URB	75	70.12	70.12	12.25	26.90	100
GDP	75	1,236	1,236	700.9	535.1	5,541
ACR	75	66.61	66.61	3.274	59.10	74.10
EMR	75	47.61	47.61	4.736	34.98	57.31
UNR	75	6.633	6.633	3.312	0.800	18.10
SHE	75	24.14	24.14	4.862	17.40	46.50
AGR	75	12.39	12.39	5.339	0.300	25.90
MAN	75	16.34	16.34	6.040	2.900	29.60
TRA	75	15.84	15.84	3.030	9.600	24.60
2008						
YOU	75	16.27	16.27	2.721	12.20	29.50
URB	75	70.10	70.10	12.31	27.10	100
GDP	75	1,320	1,320	706.5	578.8	5,496
ACR	75	67.02	67.02	3.587	56.30	75.70
EMR	75	47.85	47.85	4.824	34.89	58.44
UNR	75	7.204	7.204	3.109	0.900	18.60
SHE	75	24.50	24.50	5.014	15.80	49.90
AGR	75	11.90	11.90	5.232	0.200	25.60
MAN	75	16.10	16.10	5.918	3	29.20
TRA	75	16.13	16.13	3.150	9.600	25.40
2009						
YOU	75	16.44	16.44	2.787	12.30	30.10
URB	75	70.16	70.16	12.26	27.50	100
GDP	75	1,200	1,200	601.4	596.5	4,612
ACR	75	67.31	67.31	3.498	58.70	77.50
EMR	75	47.35	47.35	4.818	34.63	57.68
UNR	75	9.079	9.079	2.780	2.800	21.40
SHE	75	25.73	25.73	4.947	17.70	48.20
AGR	75	11.89	11.89	5.270	0.300	25.50
MAN	75	15.31	15.31	5.499	3.500	27.30
TRA	75	16.37	16.37	3.161	10	25.20
2010						
YOU	75	16.56	16.56	2.861	12.10	30.70
URB	75	70.41	70.41	12.22	27.70	100
GDP	75	1,298	1,298	656.1	592.8	4,837
ACR	75	67.43	67.43	3.281	58	78.10
EMR	75	47.53	47.53	4.863	34.45	58.73
UNR	75	8.096	8.096	2.663	1.800	21.70
SHE	75	26.39	26.39	4.942	15.80	48.50
AGR	75	11.98	11.98	5.463	0.300	25.40
MAN	75	15.21	15.21	5.464	3.100	26.60
TRA	75	16.46	16.46	3.113	10.30	25.30
2011						

YOU	75	16.82	16.82	2.927	12.40	31.40
URB	75	70.52	70.52	12.23	28.70	100
GDP	75	1,464	1,464	763.9	643.5	5,584
ACR	75	67.93	67.93	3.371	56.60	79.50
EMR	75	47.50	47.50	4.912	34.29	59.22
UNR	75	7.176	7.176	2.374	1.400	17.30
SHE	75	27.75	27.75	5.006	18	49.30
AGR	75	11.79	11.79	5.430	0.300	25.70
MAN	75	15.20	15.20	5.449	3.700	27.10
TRA	75	16.63	16.63	3.104	10.40	25.30
2012						
YOU	75	17.17	17.17	2.980	12.80	32
URB	75	70.60	70.60	12.19	28.90	100
GDP	75	1,531	1,531	772.8	670.1	5,824
ACR	75	68.07	68.07	3.505	56.50	79.10
EMR	75	47.50	47.50	4.850	33.74	58.18
UNR	75	6.145	6.145	2.436	0.800	18.40
SHE	75	29.29	29.29	4.967	22.50	50
AGR	75	11.59	11.59	5.486	0.200	26.50
MAN	75	15.08	15.08	5.466	3.900	27
TRA	75	16.77	16.77	3.097	10.30	25.20

Appendix C. Regions along axes





Appendix D. Detailed description of Statis model. Measuring disparities: three-way matrices.

As we have seen, the disparities among regions (cases) can be studied on the basis of numerous indicators (variables) like per capita GDP, productivity and the employment rate, and they can also be measured in their temporal dynamics (time). The multidimensional nature of regional differences therefore lends itself well to analysis by means of multivariate analysis methods, and in particular by dynamic multivariate analysis. We decided to apply the STATIS (Structuration des Tableaux A Trois Indices de la Statistique) (in English: Structuring Three-way data sets in Statistix) method. This is a dynamic multivariate method which enables analysis of multidimensional (multiway) phenomena expressible in the form of three-way matrices: cases i, variables j, time t. The method has been developed by Escoufier (1985), and it has found numerous applications in economics Rivadeneira et. Al 2016, in Italy as well (D'Ambra 1985; Fachin and Vichi 1993; Tassinari and Vichi 1994). Moreover, it has already been used to explain the dynamics of disparities among the Italian provinces (Amendola et al. 19979 and the European regions (Amendola et al. 2006). This technique of exploratory analysis is based on study of a three-way data matrix, $X_{I\ JT}$ obtained from the temporal succession of data matrices, ${}_{t}X_{i,j}$ of the same order, where i is the statistical unit and j the variable, both of them relative to the period t (i = 1, 2...I; j = 1, 2...J; t = 1, 2...T). The formula is:

$$X_{I,JT} = \left\| {}_{1}X \quad {}_{2}X \quad {}_{T}X \right\|$$

which can be presented as

$$X_{i,JT} = \left\| \begin{matrix} {}_1x_{11} & {}_1x_{12} & .. & {}_1x_{1j} & {}_2x_{11} & {}_2x_{12} & .. & {}_2x_{1j} & {}_tx_{11} & {}_tx_{12} & .. & {}_tx_{1j} \\ {}_1x_{21} & {}_1x_{22} & .. & {}_1x_{2j} & {}_2x_{21} & {}_2x_{22} & .. & {}_2x_{2j} & {}_tx_{21} & {}_tx_{22} & .. & {}_tx_{2j} \\ ... & ...... & .. & ...... & ...... & ...... & .. & ...... & ...... & ...... & .. & ...... \\ {}_1x_{i1} & {}_1x_{i2} & .. & {}_1x_{ij} & {}_2x_{i1} & {}_2x_{i2} & .. & {}_2x_{ij} & {}_tx_{i1} & {}_tx_{i2} & .. & {}_tx_{ij} \end{matrix} \right\| .$$

From the three-way matrix thus constructed it is possible to derive:

1. the variance-covariance matrix

$$\Sigma_{JT,JT} = \left\| \begin{matrix} {}_1\Sigma^2 & {}_{12}\Sigma & ...... & {}_{1T}\Sigma \\ {}_{12}\Sigma & {}_2\Sigma^2 & ...... & {}_{2T}\Sigma \\ ...... & ...... & ...... & {}_{pq}\Sigma \\ {}_{1T}\Sigma & {}_{2T}\Sigma & {}_{pq}\Sigma & {}_T\Sigma^2 \end{matrix} \right\|$$

where ${}_{pq}\Sigma$ is the variance-covariance matrix between *pXi,j* and *qXi,j:*

$$ {}_{pq}\Sigma = \left({}_{p}\hat{X}'_{i,j,\ q}\hat{X}_{i,j} \right)\frac{1}{n}$$

where $\hat{X}$ is the deviation matrix and 1 $1 < p < T$, $1 < q < T$

The matrices on the main diagonal represent the variance-covariance matrices of the matrix $X_{I,JT}$ at time t, while ${}_{pq}\Sigma$ measures the same relation between the variables relative to time q and time j.

2. The (TxT) square matrix $I_{T,T}$ where each generic element , $I_{p,q} = tr({}_{pq}\Sigma)$ corresponds to the trace of the relative submatrix ${}_{pq}\Sigma$ of $\Sigma_{JT,JT}$

$$I_{T,T} = \begin{Vmatrix} tr({}_{11}\Sigma^2) & tr({}_{21}\Sigma) & \ldots\ldots & tr({}_{1T}\Sigma) \\ tr({}_{12}\Sigma) & tr({}_{22}\Sigma^2) & \ldots\ldots & tr({}_{2T}\Sigma) \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & tr({}_{pq}\Sigma) \\ tr({}_{1T}\Sigma) & tr({}_{2T}\Sigma) & tr({}_{pq}\Sigma) & tr({}_{TT}\Sigma^2) \end{Vmatrix}$$

and is a measure of the dissimilarity between, *pXi,j* and *qXi,j.*. The higher the value assumed by this index, the less the similarity between the structures of *pXi,j* and *qXi,j*

Alternatively, one may assume as the index of similarity Escoufier's (1976) coefficient:

$$I^{*}_{p,q} = RV({}_pX_{i,j}, {}_qX_{i,j}) = \frac{tr({}_{pq}\Sigma_{pq}\Sigma)}{\sqrt{tr({}_p\Sigma^2)tr({}_q\Sigma^2)}}$$

obtained by operating with matrices of deviations from the mean. This is used to calculate the matrix of *RV* coefficients (*K* x *K*) called *between matrix cosine* or simply *RV matrix* and denoted by C, to analyze the similarities structure of the matrices. The *RV* coefficients are non-negative and ranges between 0 and 1, and the closer *RV* is to 1 means the more similar the two data matrices *k* and *k´* are.

The STATIS method

The STATIS method divides into three phases: *Interstructure, Compromise and Intrastructure*. The purpose of the *Interstructure* phase is to identify a suitable vectorial space smaller than T, where the T occasions can be represented. To this end, examination is made of the matrix $I_{T,T}$ (also called the interstructure matrix), the column vectors of which are assumed as characteristic elements of each of the T occasions. Constructed from this is a factorial subspace $\Re^s$ con $s<t$ generated by the s eigenvectors corresponding to the s largest eigenvalues of $I_{T,T}$ con $s<t$. The subspace thus constructed yields the best representation of the T occasions because it is demonstrated that the matrix Q, of rank s <T – whose elements $Q_{(s)} = \sum_{a=1}^{s} \delta_a u_a u'_a$ are linear combinations of the first δ_a eigenvalues and u_a eigenvectors of the matrix $I_{T,T}$ – has the characteristic of minimizing the square of the Euclidean norm $\| I\text{-}Q \|^2$.

A first result is thus obtained. The T occasions with coordinates equal to $\sqrt{\delta_1 u_1}$, $\sqrt{\delta_2 u_2}$, … $\sqrt{\delta_h u_h}$ can be generated in the factorial subspace $\Re^s$ by the first eigenvectors u_a .

It is also possible to calculate indices relative to the quality of the representation, and also relative to the contribution made by each of the T occasions:

− the ratio between the sum of the first s eigenvalues and the total of all the eigenvalues is a measure of the percentage of total information contained in the space $\Re^s$;

− the ratio between the individual eigenvalue and the overall total measures the variability captured by the relative eigenvector;

− the square of the cosine of the angle formed by the factorial axis with the segment that joins the occasion-point with the origin is an index of the representation quality of the individual occasion from that axis;

− the proximity of two occasion-points in the space $\Re^s$ is an indicator of the similarity of the matrices.

In the *compromise* phase, a fictitious structure or synthesis matrix is identified which optimally summarizes the information contained in the *T* variance and covariance matrices. This structure, called 'compromise', is given by the matrix *W* obtained as a linear combination of the elements u_1 of the eigenvector of the matrix $I_{T,T}$ corresponding to the highest eigenvector and the matrices $\Gamma_t = {}_t\hat{X}_t \hat{X}'$

$$W = \sum_{t=1}^{T} u_t \Gamma_t$$

In the space plotted by the s eigenvectors corresponding to the first s eigenvalues of the matrix *W* it is possible to represent both the j variables and the median positions of each individual. The latter are derived from the diagonalization of matrix W obtained by identifying a matrix M such that W = MM'D (where D is a diagonal matrix defined positive whose elements are the weights of the individuals, statistical units, $D = \frac{1}{L} I$, with L equal to the number of individuals, and where I is an identity matrix.

In other words, matrix W is the best compromise, in the sense defined above, among the various representations that can be associated with each of the T matrices taken separately for each unit of time.

If s = 2, the representation occurs in a two-dimensional space corresponding to the first two factors identified. Obviously, this projection will be better, the greater the incidence of the first two eigenvectors on the trace of W.

In the *intrastructure* phase it is then possible to represent the trajectories followed in time by each individual in the factorial space thus identified. If only the first two eigenvalues are considered, the representation of the trajectories may occur in a space where the system of Cartesian axes is constituted by the eigenvectors a_1 and a_2, and where the coordinates on the first axis of each individual are given by $\left(\delta_{1t}\Gamma a_1\right)^{-0.5}$ and on the second axis by $\left(\delta_{2t}\Gamma a_2\right)^{-0.5}$.