



NATIONAL RESEARCH UNIVERSITY  
HIGHER SCHOOL OF ECONOMICS

*Alexander Chepurenko, Olga Obraztsova,  
Vladimir Elakhovsky*

**CROSS-REGIONAL VARIATIONS  
IN THE LEVEL OF  
ENTREPRENEURIAL ACTIVITY IN  
RUSSIA BY TYPE OF  
MOTIVATION: DETERMINING  
FACTORS**

BASIC RESEARCH PROGRAM

WORKING PAPERS

SERIES: MANAGEMENT  
WP BRP 45/MAN/2015

*Alexander Chepurenko<sup>1</sup>, Olga Obraztsova<sup>2</sup>, Vladimir Elakhovsky<sup>3</sup>*

## **CROSS-REGIONAL VARIATIONS IN THE LEVEL OF ENTREPRENEURIAL ACTIVITY IN RUSSIA BY TYPE OF MOTIVATION: DETERMINING FACTORS<sup>4</sup>**

The paper deals with the difference in the share of opportunity-based early entrepreneurs among regions in Russia, which is an important indicator of the ‘quality’ of the entrepreneurial activity. We invent an index called the share of opportunity-based early entrepreneurs (SOBE) which is defined as the number of nascent entrepreneurs and new business owners who are driven by the search for new opportunities and towards the realization of their own values when starting-up and developing their businesses. It is shown that the differences in SOBE levels among Russian regions are statistically significant; cross-regional differences in the SOBE level reflect a certain set of regional social and economic factors right away or with an one-year or two-year lag; they may depend on the tempo of changes in a certain set of factors related to social and economic development in given regions. Among the confirmed hypotheses are the following: a successive growth of private investment in the regional economy as well as a stable increase of real wages of employed population are factors that decrease a region’s SOBE level; moreover, the higher the population’s access to PC and Internet at home, the higher is the related region’s SOBE level. The empirical part is based on the survey designed by the Higher School of Economics which was conducted in 2011 in 79 regions of Russia with a sample of 56 900 respondents. The survey is representative for the structure of the adult population in each of the surveyed regions.

**Key words:** entrepreneurial activity, entrepreneurial motivation, regional economy, entrepreneurship in Russia

**JEL:** L26 R11

---

<sup>1</sup> Prof., Higher School of Economics, Moscow, [achepurenko@hse.ru](mailto:achepurenko@hse.ru)

<sup>2</sup> Ass. Prof., Russian State Social University, Moscow, [Olga\\_Obratzsova@rambler.ru](mailto:Olga_Obratzsova@rambler.ru)

<sup>3</sup> Ass. Prof., Higher School of Economics, Moscow, [vlstel@yandex.ru](mailto:vlstel@yandex.ru)

<sup>4</sup> The paper is based on the data of the project ‘Entrepreneurship and small business: development and success factors in societies under systemic transition (2015) supported by the Center of fundamental research of the National research university – Higher School of Economics (Moscow).

## **Introduction**

The regional disparity of entrepreneurial activity, measured by varying rates of already existing SMEs as well as by the frequency of start-ups and business closures, is a widely recognized problem. There are various approaches in the literature based on labor market analysis (factors like unemployment and skills), firms' ecology (industrial structure of regional economy by size and branch), demography (population density/growth, human capital), and financial infrastructure (availability of financing, etc.) of characteristics which significantly influence regional variation in new firm birth rates.

Already in 1980-1990s some achievements were made when investigating different factors on the disparity of the entrepreneurial activity across regions (Fritsch and Storey 2014). In the middle of the 1990s, it became clear that the most evident shortcomings are the usage of only cross-sectional analysis (the result of a lack of time-series data), as well as the absence of reliable indicators to measure the impact of regional policy on the intensity of entrepreneurial start-up activity.

In 2000s, certain incremental knowledge was produced with regards to some of the questions raised earlier. However, most papers were dealing with the impact of entrepreneurial activity on varying growth and prosperity in different regions (Audretsch and Keilbach, 2004; Fritsch and Mueller, 2004; Van Stel and Storey, 2004; Mueller et al., 2008).

Lee et al. (2004) explored the role of such region-specific factors as social characteristics and human capital in the entrepreneurial performance of regions. The authors argued that social diversity and creativity have a positive relationship with new firm formation. Namely, new firms' formation is positively and significantly associated with social diversity and insignificantly with creativity. Hence, it is necessary to pay attention to the 'social habitat of a region' to understand its regional entrepreneurial dynamics. However, it is still unclear whether the higher entrepreneurial activity predicts the level of human and cultural capital and economic development or is a result of them.

As Audretsch, Bonte & Keilbach (2008) discovered using data from 310 West German counties, innovation efforts have an indirect effect on economic performance via entrepreneurship. On the one hand, innovation efforts of start-ups lead to an increase in regional technical knowledge, improving the regional industry's economic performance. On the other hand, regional innovation efforts also increase entrepreneurship capital which, in turn, improves regional economic performance.

Knowledge and, again, cultural diversity as factors of regional differences in entrepreneurial activity remain significant in the literature – for example, Audretsch, Dohse & Niebuhr (2010) investigated the determinants of entrepreneurial activity in German regions in 1998-2005. Their regression analysis showed that regions with a high level of knowledge provide more opportunities for entrepreneurship than other regions. Moreover, cultural diversity has a positive impact on technology oriented start-ups. Thus, regions characterized by a high level of knowledge and cultural diversity form an ideal breeding ground for technology oriented start-ups.

Moreover, as is pointed out by Fritsch and Storey, ‘although several studies find a statistical relationship between personality traits and personal attitudes of the regional population and the level of entrepreneurship, they are unable to identify the causality of these effects. Do specific value-sets amongst the population of a region bring about relatively high or low levels of entrepreneurship or is it entrepreneurship that causes the expression of these values?’ (Fritsch and Storey, 2014, p.945).

These person-related variables have to do with individual characteristics, also called perceptual variables (opportunity cognition, self-efficacy, fear of failure), education level and practical experience, prior employment in a small firm or self-employment experience (Zhao and Seibert, 2006; Rauch and Frese, 2007; Zhao et al., 2010; Storey and Greene, 2011; Sorgner and Fritsch, 2013; Caliendo et al., 2014). For instance, regions with a high share of people inclined to become entrepreneurial (Obschonka et al., 2013) or employees of small firms could have higher entrepreneurial entry rates. However, such explanations are subject to endogeneity danger as individual behaviour can be affected by the regional institutions.

In any case, to deal with personal variables, a regionally representative sample of entrepreneurs is needed. Unfortunately, such data cannot be obtained either from the GEM or from the PSED as these projects are run on the basis of a nationally representative sample and are not fit to cover the region representative samples.

Moreover, longitudinal historical observations, like a study by Fritsch and Wyrwich (2013) reflecting the culturally embedded differences in entrepreneurial perceptions in different parts of Germany, are questionable as there is a lack of reliable statistics at the regional level.

Two more approaches are important for the present paper. First, Bergmann (2004), using multivariate analysis methods data for 10 German regions, showed that regional start-up activities, when determined by demand and agglomeration, depend in a more direct way on attitudes and self-efficacy of novice entrepreneurs which mediate the objective state of the economy and infrastructure and other framework conditions of any region. Similar evidence was

found by Bosma and Schutjens (2011) who compared the data for 127 regions in 17 European GEM participating countries, linking institutional factors and economic and demographic attributes to variations in regional entrepreneurial attitude and activity. In particular, entrepreneurial attitudes (the fear of failure in starting a business, perceptions on start-up opportunities in the settlement of residence and self-assessment of personal capabilities to start a firm are among them) influence differences in TEA and related indicators. But what, in turn, are the factors determining the differences in motivation across regions, and in particular in such specific environments like transitional economies?

More recently, Kibler et al. (2014) paper showed the role of the differing levels of social legitimacy of entrepreneurship in different regions affecting the formation of new businesses. The authors stress that a better attitude toward entrepreneurship 'can give those with entrepreneurial intentions the final impulse needed to turn their intentions into actual start-up behaviour'. Moreover, Bosma and Sternberg (2014) argue that urban environments are more prospective for opportunity-driven, or opportunity-based entrepreneurship.

There are few papers on regional differences in entrepreneurial activity in transitional and developing countries. Naude et al. (2008), however, used data from South Africa to empirically identify the determinants of start-up rates across different regions, concluding that the most important determinants of start-up rates across South Africa's districts are profit rates, educational levels, agglomeration (measured by the economic size of a district), and access to formal bank finance. The second important finding is the insignificance of unemployment as a 'push' motive to start-up. The authors also have found that access to formal bank finance matters for regional start-up rates, which is not typical for a developing country and that market-size (agglomerations) is negatively associated with start-up rates in South Africa which may imply that there are some other factors in bigger settlements such as higher competition and/or economic barriers to entry (monopolistic behavior, etc.).

Unfortunately, the difference of the entrepreneurial motivation structure among regions and its reasons are less investigated. And with regard to transitional economies like Russia, the present paper is the first attempt to this theme.

The reasons of such a state are evident: a lack of reliable statistics on early entrepreneurial activity (which is usually informal/non-observed) at the regional level. Even such projects as the RLMS which is very similar to the British Household Panel Survey, or the German Socio-Economic Panel, do not provide enough data to make cross-regional comparisons. Therefore, we used the data of the pilot project conducted by means of the 'Georating' survey of the 'Public Opinion' Foundation (FOM) in 2011.

This paper deals with the problem of spatial heterogeneity of entrepreneurial motivation in Russia and its factors<sup>1</sup>. It is structured as follows. In the beginning section, the approach and the methodology of the research are described. Here, first, the source of the data and the design of the survey are explained. Second, as the rather seldom used in the entrepreneurship research FLDA technique is applied, a comprehensive argumentation of the methodology is given. Then, the findings, i.e., the set of factors of cross-regional diversity of entrepreneurial motivation in Russia, are explained; finally, the evidence and political recommendations, as well as the limitations of the research are discussed.

### **Hypotheses, data and methodology**

As the analysis of the available literature shows, the variation in the entrepreneurial motivation structure among regions and its reasons not widely investigated. And regarding transitional economies such as Russia, the present paper is the first attempt to discover them.

There are some constraints to such an investigation in Russia to be taken into consideration. First, the impact of already existing SME on both national and regional economy should be very moderate, as it does not play any vital role in economic growth; according to the Russian Statistics Agency (Rosstat) data from April 2014, the SME sector is providing ca. 20 % of GDP and ca. 23 % of employment in Russia. Second, the fiscal system in Russia is structured ‘from top to the bottom’; state tax incomes from entrepreneurship (income tax for solo owners and profit tax for firms) are not only very limited, but also the biggest part of them is filling the federal, not the regional and local budgets. Hence, even a relatively high density of already established SME is neither providing financial resources nor shaping incentives for regional or municipal authorities to promote entrepreneurial activity. Being co-funded by the Federal state, respective SME support programs are structured similarly in heterogeneous regions and locations (Chepureenko, 2011). Therefore, the paper does not accept the role of the policy of regional authorities as a factor influencing the entrepreneurial activity or its motivation structure (Fritsch and Mueller, 2007; Caliendo and Kuenn, 2014).

In the literature, especially after the GEM data became available for such kind of research, it is widely accepted that the quality of the entrepreneurial activity matters; the impact of opportunity-driven entrepreneurs on economic growth and wellbeing is higher than the impact of necessity-driven entrepreneurs (van Stel et al., 2005; Wennekers et al., 2005; Wong et al., 2005;

Hessels et al., 2008; Koellinger, 2008; Valliere and Peterson, 2009; Hessels et al., 2010; Chepurensko et al., 2011). As in any location there are different types of entrepreneurs presented; some of them are pushed to become entrepreneurs, while some others by contrast are rather pulled by the opportunity. As a person's motivation is predicted by a set of various economic, social, institutional and societal specificities of the given region, the structure of the entrepreneurial activity by motivation is a unique picture reflecting the complex of different factors positively or negatively affecting entrepreneurial activity in these local environments. Thus, the higher level of opportunity-driven entrepreneurship is an important signal of the possible perspectives of the respective region, and vice versa.

The share of opportunity-based early entrepreneurs (SOBE) for this analysis was defined as the number of nascent entrepreneurs and new business owners who according to their answers were driven by the search for new opportunities and towards the realization of their own values when starting-up and developing their businesses.

The following hypotheses have been determined related to regionally specific factors which act either as stimuli for opportunity-based motivation or as barriers.

As is shown in the previous literature, the motivation to become entrepreneurial is in many ways dependent on macro-level factors. The statistics of socio-economic data as well as the available data concerning the current density of SMEs and the entrepreneurial activity of population in different regions of Russia show a high level of variation. It seems that such differences may result in a big dispersion among Russian regions in the structure of the motivation of early entrepreneurs. Hence:

***H1. The differences in SOBE levels between specific groups of Russian regions are statistically significant.***

It is evident that the improvement or worsening of the economic situation of households, changes in social welfare of the population, etc., do not work as factors which push or pull people to start a new venture immediately; the perception of the new situation, decision making, searching for opportunities and resources as well as the actual start-up of a business takes a certain amount of time. Moreover, the dynamics of the changes of certain socio-economic indicators might be important, too, because humans' reaction may vary according to the tempo of the changes: a more dynamic improvement or deterioration of living conditions, labor market situation, etc., may foster or weaken their willingness to start a business. Based on these assumptions, the second and third hypotheses are formulated.

***H2. Cross-regional differences in the SOBE level may reflect a certain set of regional social and economic factors with an one-year or two-year lag.***

*H3. Cross-regional differences in the SOBE level depend on the tempo of changes in a certain set of factors related to social and economic development in given regions.*

As the opportunity-based entrepreneurial activity is in many ways a result of a rational choice between starting one's own business and employment based on the estimation of future wins and losses (opportunity costs), there should exist some factors which could diminish the willingness of adults in respective regions to become entrepreneurial (a certain 'free rider problem'), like a growing economic dynamic and increasing benefits of employees from it. In such a situation, the willingness to discover and use entrepreneurial opportunities should diminish. Based on this observation, the following hypotheses are formulated:

*H 4.1. A growth of private investment in the regional economy could become a factor that decreases a region's SOBE level.*

*H 4.2. An increase of wages of the employed population could become a factor that decreases the respective region's SOBE level.*

Opportunity-based entrepreneurship has to do with the rational search for exploration and exploitation of entrepreneurial opportunities. Information technologies contribute to store, manipulate, distribute or create information as well as to form and develop social networks, so they might be useful for the exploration of entrepreneurial opportunities by people who are obtaining a higher human capital (education, skills, etc.). Meanwhile, such people are more inclined toward an opportunity-based type of entrepreneurial motivation.

*H 5. Cross-regional diversity in the access of households to PC and the Internet at home, determine differences in the SOBE level. The higher the access, the higher is the SOBE level of the related region.*

To prove these hypotheses, the data of the pilot project conducted by means of the 'GeoRating' survey of the 'Public Opinion' Foundation (FOM) in 2011 were used.

### ***Quality of data and survey methodology***

To estimate the regional differentiation in entrepreneurial activity of Russia's population the main principles of the explanatory model and APS Methodology of the Global Entrepreneurship Monitor (see Reynolds et al., 2005) are used; but the size of the sample needs to be large enough to represent both the overall population and the population structure of all 79 surveyed Russian regions to measure the entrepreneurial activity in the respective regions.

When mastering the sample, a 3-stage stratified household sampling procedure was used. At the first stage, the overall Russian population was clustered by geography, and administrative

districts were used as primary sampling units (PSU). PSUs represent 79 (of 82) Russian regions with 139.9 million inhabitants, or 98.6% of the overall population of the Russian Federation. The PSU sample size varied from 500 respondents in 21 regions to 800 respondents in 58 regions, depending on the size of the population. The total sample size was 56.900 adults, from 18 years and older.

Specific locations (2.335 in total) were used as secondary sampling units (SSU). SSUs included 77 administrative centers (Moscow and St. Petersburg among them), 683 cities, 282 small towns (official status – urban-type settlement) and 1.293 villages. There were 2.372 cities and urban-type settlements in the selected 79 regions, and 1.042 of them (i.e. more than 1/3) were included in the sample.

At the third stage observation units (households) were selected by Random Route Sampling Method. Selection of a respondent in a household was made on the basis of a connected quota by sex and age and a separate quota for education.

Sampling error did not exceed 5.5% for regions with 500 respondents and 4.6% for regions with 800 respondents. The total statistical error for the overall Russian population did not exceed 1%.

The survey was conducted in May 2011.

Such an approach ensured the database harmonized with the GEM method of entrepreneurial activity observation.

The questionnaire consisted of only 18 questions, eight of them were related to socio-demographic characteristics (gender, age, education, professional occupation and status, wellbeing) and 10 referred to issues of entrepreneurial potential and the activity of the population in accordance with the GEM Methodology (Obraztsova and Popovskaya, 2012). These questions enabled a differentiation in entrepreneurial cohorts, the type of entrepreneurial motivation, the sources of financing, and the respondent's opinion on entrepreneurial framework conditions in the area where he/she lives.

According to the GEM methodology, there are different entrepreneurial cohorts defined by the stage of their activity; namely, there are potential entrepreneurs still only expecting to start in nearest future; novice entrepreneurs, or baby business owners (BBO) whose entrepreneurial experience does not exceed 3 months; new business owners (NBO, from 3 until 42 months); and, lastly, established business owners managing ventures older than 42 months. Two cohorts, BBO and NBO, form a group called early entrepreneurs. The share of early entrepreneurs, or the total entrepreneurial activity index (TEA) which is the theme of the present paper.

## **The analysis of variation of the regional share of opportunity driven early entrepreneurship (SOBE)**

To develop cross-regional comparisons of the SOBE level in Russia, a statistical analysis was made of regional distribution by this indicator observed in 2011 (on the base of descriptive statistics of variations - *see Table A1.1 in App.1*).

All regions were divided in homogeneous groups by SOBE level. Among them, 3 regions in the Central Federal District of Russia are characterised with zero level of SOBE. Hence, they were excluded from the cluster analysis, and the number of homogeneous groups of the regions with non-zero level of SOBE was determined according to Sturgis's criteria, using k-means cluster analysis to identify various clusters on the base of SOBE in 2011 (*see Table A1.2 in App.1*). The composition of the resulting groups was then optimized through an iterative process of determining that k value which would yield a step-like increase in the maximum among-group variation (sum of squares among groups - SSA) of the  $\sigma^2_{SSA}$  value, going from minimum to maximum values (on aggregate). The result was a stable 5-cluster structure (*see histograms in Figure 1 below*) homogeneous by non-zero SOBE level (labelled as low, below than average, average, above average or high). The average level is not significantly lower than the average SOBE level in GEM countries in 2011 (GEM Global Report 2011, 2012).

The standard variation analysis was used to study those regions' SOBE distributions. The evaluation of the significance of these SOBE level differences' among groups of regions was estimated on the basis of Student's t-criteria for checking the first hypothesis formulated above.

### ***Factors of the regional SOBE variation: variables and methodology***

As the second step, a multi-dimensional analysis of factors influencing the motivation to start a business was mastered. Non-parametric scales, coefficients and methods were used as the regional distribution significantly by SOBE level differs from a normal distribution (t-criteria with p-value 0.005). The dependent variable was an ordinary label of a group of regions homogeneous by SOBE level. To explore which regional social and economic variables could influence early stage entrepreneurs' motivation, statistical data of the Russian Federal Statistical Service's on 123 indicators of the economic and social situation in Russian regions and their dynamics for 2008-2011 were used. The broad range of indicators (*see the list of variables in App.2*) included several characteristics of the demography, labour market, wellbeing and

poverty, economic and especially investment activity, the state of the physical infrastructure, the level of crime. For the reasons mentioned earlier, these indicators were taken both for the same year and with a lag of one or two years. We used either level or tempo (including with a one-year or two-year gap) of these indicators. The relationship between SOBE level and regional external factors was measured on the base of Spearman's Rho coefficient. Then, only indicators that had a statistically significant relationship with the regional SOBE level at a confidence level of 5% were been tested as independent variables in a model of factor analysis.

As the dependent variable was measured on an ordinary scale, Fisher's Linear Discriminant Model was applied to prove which special regional factors enhance the regions' group recognition.

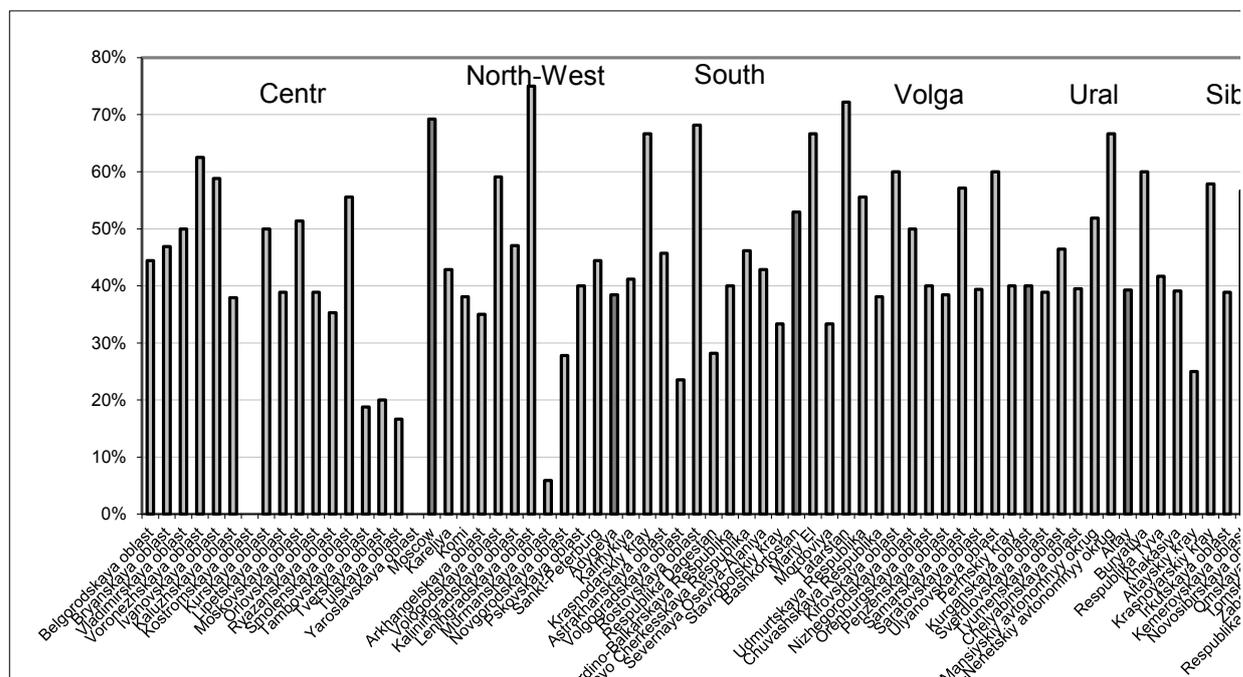
It was necessary to compare several characteristics of Russian regions with high or low or zero SOBE level and to prove the correlation significance between regional peculiarities and the level of SOBE. An appropriate statistical technique to examine whether two or more mutually exclusive groups of territories can be distinguished from each other, based on linear combinations of values of independent predictor variables and to determine which variables contribute to the separation, is the linear discriminant analysis (LDA) (Fisher, 1936; McLachlan, 2004). The LDA model can be used on the basis of information about the reckoning of certain Russian regions to a respective group and about the annual dynamics of the economic and social indicators with a short term gap for each of them, but not the canonical LDA Model. Fisher's linear discriminant analysis (FLDA) was applied because, contrary to the canonical LDA, this method do not use hard assumptions concerning the statistical characteristics of the initial data; it does not necessary imply a normal distribution and quantitative scales for the variables. Besides, it prefers attributive measurement scales (Lim et al., 2000). FLDA is used in this research to find a linear combination of regional features which characterize or separate for low, average and high SOBE level groups of Russian regions.

The list of 24 independent variables was built as a result of plotting pairs of independent variables to check whether the relationships among them are approximately linear. At the next stage of analysis the variables which maximize Mahalanobis' distance between the centroids of groups were selected for entry into the equation of the discriminant function.

Then, the FLDA model was the statistical instrument to find factors determining cross-regional differences observed in Russia in 2011.

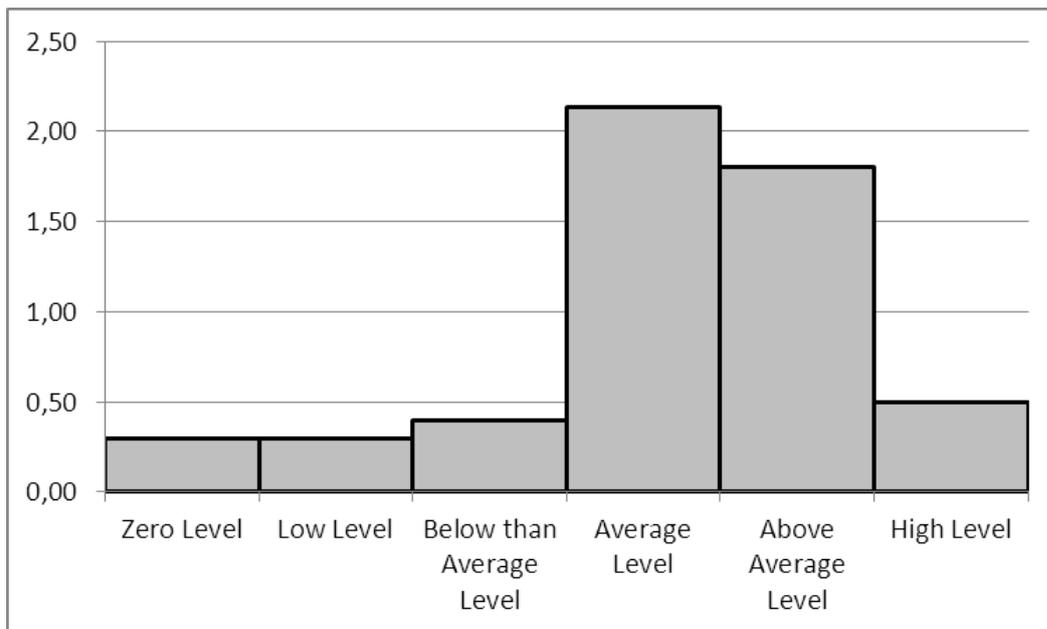
### **Opportunity-based motivation of early entrepreneurs in Russian regions: findings**

The analysis shows that there is a high dispersion of regions as regards the SOBE levels (see Figure 1); the range of the SOBE distribution is 75%. Moreover, there are statistically significant differences between regional levels of the indicator within the same space boundaries that may be described on the basis of federal districts (Student's t-criteria with a 1% confidence level).



**Fig.1.** Regional SOBE levels in Russian Federation (clustering in space boundaries, in alphabetical order)

Cross-regional comparisons show that regional SOBE indicators demonstrate a significant level of variation (with a coefficient of variation near 90% and half-quartile range more than 30% – see Tables A1.1 and A1.2 in App.1), while the average SOBE value was around 28% in Russia in 2011 according to both the GeoRating and the GEM data. Using the GeoRating data the distribution of Russian regions by SOBE level in 2011 was constructed and analysed (see Figure 2). The population of Russian regions is not homogeneous by SOBE level; the coefficient of variation is higher than 35%. According to the cluster analysis, Russian regions were divided by SOBE levels into five homogeneous groups (Figure 2). The left tail of the distribution is longer; most of the regions are concentrated on the right of the figure (Lindberg's skew coefficient is -0.27). Moreover, the distribution demonstrates a significant peak and the absence of the right shoulder (Lindberg's excess coefficient is +0.123).



**Fig. 2.** Homogeneous groups of Russian regions' by the SOBE level (the ordinate indicates density of distribution)

The GEM average level of the SOBE is the most typical for Russian regions, and this cluster consists of 32 territorial entities. The clusters of regions with the SOBE levels at the GEM average and above dominate in the Russian Federation. Both the zero and low level clusters consist of only 3 regions, while the cluster with the SOBE level below average includes only 4 regions (Table A1.2), but these regions form a significantly heavy shoulder of regional distribution taking down average level of the SOBE in Russia.

The index of qualitative variation (IQV) is a measure of the variability for categorical variables. It is based on the ratio of the total number of differences in the distribution to the maximum number of possible differences within the same distribution and takes a value from 0 to 1; if all the cases belong to the same category there is no variability, and the IQV is equal zero. In contrast, when the cases are distributed uniformly across the categories, there is maximum variability and IQV is 1 (for more detail, see Frankfort-Nachmias and Leon-Guerrero, 2014, pp. 138-140). IQV for the groups of regions were calculated to evaluate the differences in the context of their space homogeneity. These IQV varied from minimum (for SOBE low-level group of regions) to maximum. The analysis showed that the clusters of regions that are homogeneous by the SOBE level significantly differ from each other, however, as a rule there is no space homogeneity within these clusters (Table 1).

**Table 1.** Significant distances between final cluster centers (SOBE) and levels of space homogeneity within clusters

Cluster	Zero Level	Low Level	Below Average Level	Average Level	Above Average Level	High Level
Zero Level	-	16,5	24,1	37,8	50,3	60,1
Low Level	16,5	-	7,6	21,3	33,8	43,6
Below Average Level	24,1	7,6	-	13,7	26,2	36,0
Average Level	37,8	21,3	13,7	-	12,5	22,3
Above Average Level	50,3	33,8	26,2	12,5	-	9,8
High Level	60,1	43,6	36,0	22,3	9,8	-
Final Cluster Centers						
Cluster	Zero Level	Low Level	Below Average Level	Average Level	Above Average Level	High Level
SOBE	2,00	18,5	26,1	39,8	52,3	62,1
IQV	0,889	0,00	1,00	0,984	0,944	0,984

The distances between the final regional cluster centres are statistically significant (t-criteria with p-value 0.005 for all inter-cluster distances except the low and below the average level clusters when p-value is 0.05). The decile dispersion ratio which measures the differentiation between the regional SOBE levels exceeds 3. The semi-quartile range of 0,568 is a good measure of spread to use for skewed distributions; it is hardly affected by the higher values of the SOBE. It means that the differentiation of Russian regions by the SOBE values is statistically significant.

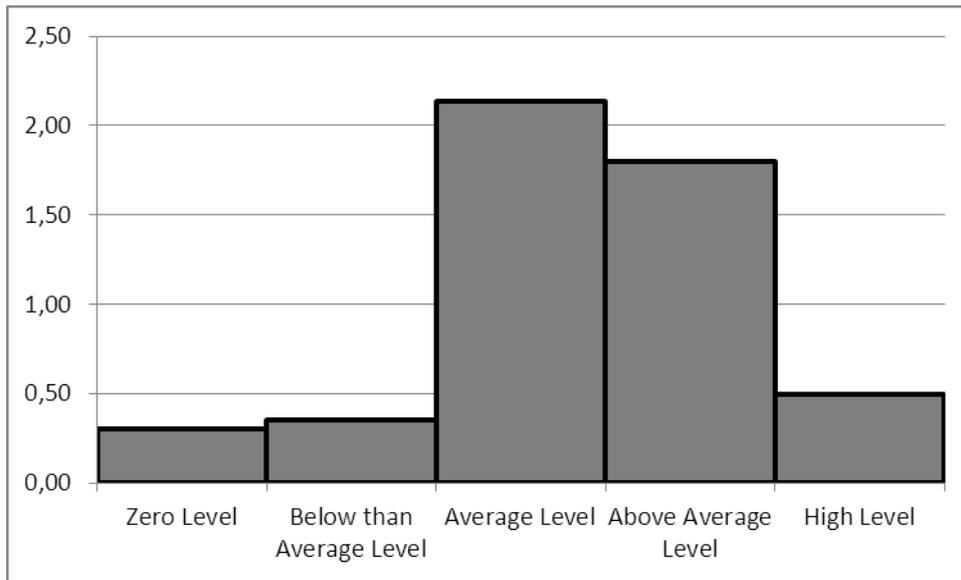
The results of the detailed analysis, at first glance, suggest that there is no direct relationship between levels of overall economic and social development or economic and geographical location and regional SOBE level.

### **Factors determining cross-regional differences**

To compare the impact of several factors of socio-economic development of Russian regions with different SOBE levels and to prove the significance of the correlation between them and SOBE, a statistical technique based on combinations of the values of the predictor variables was used to examine whether several groups of regions can be separated. Then, it was checked using the FLDA to see which variables contribute to such a separation.

The homogeneous groups of regions were constructed according to the SOBE to get a relevant clustering of the whole sample. The groups with low and below average level of the SOBE were

integrated into one group to increase the significance of a distance between these clusters centres in order to improve the quality of classification in FLDA model (Figure 3).



**Fig. 3.** Homogeneous groups of Russian regions by the SOBE level (the ordinate indicates density of distribution; low level and below the average level groups are integrated)

The level of the homogeneous group of regions was a dependent variable in the FLDA procedure whilst the set of socio-economic indicators served as a combination of independent variables (see Appendix 2).

To emphasize again, it was suggested that the SOBE levels are dependent not only (if not primarily) on the set of actual social and economic variables, but also from their dynamic over time.

Thus, on the basis of FLDA procedure a subset of variables was identified that do not meet the removal criterion and can be applied to separate the groups of regions. After all of the preliminaries, the independents were entered to compute the coefficients of the discriminant function and to calculate the discriminant scores which are combinations of the predictor variables (Table 2).

It is to emphasize that the cyphers of the coefficients are arbitrary. Negative coefficients could just as well be positive if the signs of the positive coefficients were made negative. One might have to look at coefficients with the same sign to determine how the variables relate to the groups. All predictors in the table are divided into 2 groups – stimuli and barriers.

**Table 2.** Canonical coefficients for significant discriminant functions

	Stimuli			Barriers	
S1	*The share of households had an actual	0,410	B1	*Investments in fixed capital per capita _	-0,274

	Internet access_2011			tempo_2010	
S2	*The share of households had a personal computer_2011	0,364	B2	*Average monthly nominal wages of workers, tempo_2010	-0,266
S3	*Average monthly nominal wages of workers, tempo_2011 <sup>a</sup>	0,312	B3	*The average annual number of employed in the economy_2010	-0,218
S4	The number of small businesses per 10000 inhabitants_2010	0,249	B4	*The average annual number of employed in the economy_2011	-0,218
S5	*The average size of the deposits of physical persons in foreign currency in Sberbank,_tempo_2010	0,206	B5	*Living space per capita (at the end of the year, square m)_2010	-0,214
S6	*The unemployment rate, %_2011	0,194	B6	*Living space per capita (at the end of the year, square m)__2009	-0,213
S7	*Number of recorded crimes per 10000 inhabitants _tempo_2010	0,193	B7	Living space per capita (at the end of the year, square m)__2011	-0,213
S8	Number of families, registered to need dwelling, %_2009	0,166	B8	The ratio of average income per subsistence level_2009	-0,147
S9	The unemployment rate, %_2010	0,152	B9	Population density, pers. per 1 square km _2010	-0,134
S10	*Number of families, registered to need dwelling %_2010	0,147	B10	*Population density, pers. per 1 square km _2011	-0,134
S11	The unemployment rate, %_2009	0,14	B11	*Population density, pers. per 1 square km _2009	-0,134
S12	*The unemployment rate _tempo_2010	0,12	B12	*Retail trade turnover per capita, thous. RUR_2011	-0,124
			B13	*Number of families, registered to need dwelling, %_2011	-0,12

\* variables associated with the first function

The variables in Table 2 are sorted and numbered according to the decreasing of the first and second discriminant functions' coefficients and divided in two sets of factors. Among those influencing it positively with only a one-year-lag there are:

- The increase of the share of households with dwelling problems;
  - the increase of the registered SMEs in the respective region;
  - the increasing of the unemployment rate;
  - the increasing average size of the bank deposits in foreign currency in Sberbank;
  - the number of recorded crimes.

The only factor with a positive two-years-lag impact on the SOBE level in the region is the unemployment rate: the higher is the unemployment rate in a region, the higher will be the SOBE level in two years; the unemployment rate is a leading indicator of the SOBE.

Factors with a negative influence on the SOBE level in the region with a one-year-lag are:

- the positive dynamics of the investment in fixed capital;
- the growing ratio of the average income to the official subsistence level.

Factors which show a more prolonged but negative influence (two-years-lag) are:

- the living space per capita;
- the share of households with dwelling problems (negative impact, i.e. an increase of this share bring down a SOBE level in the given region).

Moreover, a single factor was found which influences the SOBE level in the same year and with a one-year-lag but in opposite ways, namely the average monthly nominal wages of employees. Its growth in the same year is positively related with the SOBE level in the respective region, but negatively with a one-year-lag.

In the list of the stimuli (Table 2) 8 predictors are level variables, and 4 are tempo indicators. There are also 2 tempo predictors decreasing the SOBE level (B1 and B2), namely, the increase of investments in fixed capital per capita and the average monthly nominal wages of employees with a one-year lag. This means that both the growing investment activities of bigger firms as well as the increasing salaries diminish the pull factors for adults in the respective region to start a venture in the coming year.

Furthermore, the share of households with internet access and the share of households with a PC are significant factors of the increasing of the SOBE level in a respective region. The successive growth of private investment in the regional economy as well as the increase of real salaries play the role of barriers for opportunity based entrepreneurship in the region.

At the final stage of the exploration of data, the independents were entered to calculate discriminant scores that are linear combinations of the predictor variables (see Table 3). Within five groups, there are four discriminant functions.

**Table 3.** Discriminant functions' eigenvalues and dispersion explained

Function	Own value	% of dispersion explained	Cumulative %	Canonic correlation
1	6,282 <sup>a</sup>	68,3	68,3	,920
2	3,226 <sup>a</sup>	17,7	86,0	,831
3	1,165 <sup>a</sup>	8,9	94,9	,755
4	1,017 <sup>a</sup>	5,1	100,0	,674

a. In the analysis first 4 canonic discriminant functions are used

The eigenvalues for each of the discriminant functions reflect how strongly the functions are related to the groups, as an eigenvalue is the ratio of the between-groups to the within-groups sum of squares for the discriminant function scores. The objective was to find a linear combination of values of the independent variables (see Table 2) which separates different groups of regions, homogeneous by the SOBE level in the best way. The functions go from best to worst, in terms of the ratios of the between-groups to within-groups sums of squares. Of 4 functions derived, not all should contribute to the separation of the groups. In Table 4, Wilks' Lambda and Chi-square show a high statistical significance of the first function and a sufficient significance of the second one.

**Table 4.** Discriminant functions' significance testing

Testing of functions	Wilks' Lambda	Chi-square	Degrees of freedom	Significance
from 1 to 4	0,004	325,831	196	0,002
from 2 to 4	0,027	145,291	144	0,017
from 3 to 4	0,208	71,126	94	0,684
4	0,602	26,119	46	0,851

These functions were used to distinguish clusters of Russian regions as aggregate latent factors explaining the SOBE levels. The scores were evaluated using the non-standardized discriminant function coefficients in Table 2. The scores can be understood as predicting factors specifying the entry of a certain region into a certain group. The number of cases which are correctly classified (compared to the original clusters) on the basis of the discriminant functions is 90.6%. So the aggregate factors mentioned above are highly significant for the differences by the SOBE level which is an indicator of the structure of early entrepreneurs' motivation between the Russian regions. The structure matrix obtained as a result of FLDA procedure (see App. 3) may help in interpreting these factors.

*Factor 1:* In Russia, a worsening of the socio-economic situation (the criminal situation, depopulation and a contraction of the effective demand) in a region with a large amount of savings of population increases the SOBE level with a one- or two-year lag when the level of the Internet penetration is high and if the possibility to invest in real estate for households is low.

*Factor 2:* In Russia, a worsening of the socio-economic situation (the criminal situation, depopulation and a contraction of the effective demand) can first reduce and then increase the SOBE level with two-year lag in a region with a higher level of the SME sector density, if the

possibility to invest in real estate for households (as an alternative for entrepreneurial activity) is low.

### **Conclusions, policy recommendations, and constraints of the research**

The data analysis has shown that there are some significant predictors of the SOBE and of the differences in SOBE levels between Russian regions. Such an exploration of factors which influence cross-regional differentiation in the structure of entrepreneurial motivation was not been carried out in other large economies of the world.

We show that differences in the SOBE level between certain groups of Russian regions are statistically significant. Hence, H1 is confirmed. This evidence is important, as it may strengthen the argument that different groups of regions in Russia need different sets of political measures to improve or foster opportunity-driven entrepreneurial activity; for this reason, specific target groups should be selected and relevant measures of support should be formulated for each type of regions, starting with those with only necessity-driven entrepreneurship and up to regions with high prevalence of opportunity-driven entrepreneurship.

Moreover, a set of regional social and economic factors influence the SOBE level of the regions with a one-year or two-year lag, either positively or negatively; the SOBE level in a region may depend also on the tempo of the changes in regional social and economic indicators (see Table 5). Hence, the H2 and H3 are confirmed, too.

Knowing the set of these factors and their impact on the SOBE in the region is a good precondition for adjusting the forecast of the future state of entrepreneurial activity in the respective region.

**Table 5.** Set of factors (stimuli + and barriers -) significantly influencing the SOBE level in Russian regions\*\*\*

<b>Factors** (S - Stimuli; B – Barriers) influencing SOBE:</b>	<b>without lag</b>	<b>with 1-year lag</b>	<b>with 2-year lag</b>
<b>Information Society</b>			
The share of households with Internet access	+ (level) S1*		
The share of households with a PC	+ (level) S2*		
<b>Economics</b>			
The number of small businesses per 10,000 inhabitants		+ (level) S4	
Investments in fixed capital per capita		— (tempo) B1*	
Retail trade turnover per capita, RUR 1000	— (level) B12		
<b>Labour Market</b>			

Unemployment rate	+ (level) S6*		
Average monthly nominal wages of employees	+ (tempo) S3*	— (tempo) B2*	
Unemployment rate		+ (level) S9, +(tempo) S12	+ (level) S11
The average annual number of employed in the regional economy	— (level) B4*	— (level) B3*	
<b>Social and Demographic Factors</b>			
The average size of the deposits in foreign currency in Sberbank		+ (tempo) S5*	
Number of recorded crimes per 10,000 inhabitants		+ (tempo) S7	
Ratio of average income to subsistence level			— (level) B8
Population density, pers. per 1 square km	— (level) B10	— (level) B9	— (level) B11
Number of families, registered as needing a dwelling	— (level) B13	+ (level) S10	+ (level) S8
Living space per capita (at the end of the year, square m)	— (level) B7	— (level) B5*	— (level) B6*

\* This variable is associated with the first discriminant function

\*\* Stimuli and barriers are numbered in descending order of significance for the SOBE level.

\*\*\* Tempo indicators are marked by ‘+’ if an increase of this independent variable correlates with the highest SOBE levels and by ‘—’ if a decrease of the independent variable correlates with the highest levels of SOBE (in the same year, and with a 1 or 2 year lag). Level indicators are marked by ‘+’ if a high value of this independent variable correlate with highest SOBE levels and by ‘—’ if it were otherwise (in the same year, and with a 1- or 2-year lag).

Among such factors, there are investments in fixed capital per capita and an increase of nominal wages. It came out that the SOBE level diminishes, if the investment in fixed capital per capita in the respective region is growing in the previous year. As the increase in the rate of investment in fixed assets may become significant when it is combined with the expansion of large businesses, persons with higher human and social capital may choose a better paid employment in such a big businesses than establishing of a new venture. Thus, H 4.1 is confirmed.

As regards the increase of nominal wages, the situation is more complicated. When the tempo of wage increases in a region is growing, the SOBE level in the same year is growing too, maybe because the employees recognize better chances to get their start-ups financed either from their own incomes or having more attractive conditions when accessing third parties financing. But the impact of the tempo of wage increases in a region on the SOBE level with a one-year lag is negative. Hence, when the tempo of wage increases in the previous year, people may become more skeptical regarding their own entrepreneurial activity, as they suppose that the ‘economic engine’ of the regional economy is working quite well and also because those who are more

entrepreneurially motivated, already started in the previous year. Thus, the H 4.2 is confirmed for a short-term (one-year) lag.

These findings are quite unexpected. The simplistic view that any advantages in economic activity in the region foster chances for a growing opportunity-driven entrepreneurial activity should be at least proved using data of other transition economies and large economies as well.

H5 is confirmed; possessing over a PC and a stable internet access, indeed, is a factor increasing the SOBE levels in Russian regions. A direct and strong correlation between the development of the ‘information society’ and opportunity-driven entrepreneurial activity in a region is evident. Hence, the policy makers on the regional level should be confident that indirect support of entrepreneurship such as growing IT-literacy, widening of broad-band internet access and diminishing the digital divide may enhance the chances for grass-roots opportunity-driven entrepreneurship even more significantly than simple providing small scale start-up funding.

It seems that in Russian regions with an initially large amount of financial savings and the highest internet penetration and high real estate prices, such as both capitals Moscow and St. Petersburg, a worsening of the socio-economic situation, which has began since the 2014-2015, may improve the structure of entrepreneurial activity in favor of opportunity-based entrepreneurs, however, with a one- or two-year lag. But the regional governments should be advised to secure the predictability and transparency of entrepreneurship support to ensure opportunity-motivated adults to start-up in the region where they reside.

Furthermore, a worsening of the socio-economic situation (criminal situation, depopulation and shortening of the effective demand) can first reduce and then increase the SOBE level with a two-year lag in a region with a relatively high level of development of the SME sector density, if the possibility to investment in real estate for households is rather low. This, we assume, is the nearest future of a few of Russian regions with the number of SME higher than average, especially some Northern Caucasian republics, where authorities have to invent some stimulating measures to prevent those adults who are ambitious, educated and entrepreneurially motivated, from leaving the region. For this reason, they should find a sound balance between the general measures of the active labour policy and a focused approach to support this segment of most prospective new ventures.

The most important constraint of the present research is the uniqueness of the data. It is difficult to assert unambiguously whether the findings above are specific to Russia or they have a more general value. It would be useful to repeat the survey or to analyze data of some other large economies with many different regions (especially of some economies in transition like Kazakhstan or China). Such analysis would check the validity of our results in a general context.

## *Literature*

- Acs Z., Desai S., Klapper L., 2008. What does “Entrepreneurship” data really show? *Small Business Economics*. Vol. 31 (3), pp. 265-281.
- Audretsch, D.B., Boente, W., Keilbach, M., 2008. Regional entrepreneurship capital and its impact on knowledge diffusion and economic performance. *Journal of Business Venturing*. Vol. 23, pp. 687-98.
- Audretsch D.B., Dohse D., Niebuhr A., 2010. Cultural diversity and entrepreneurship: a regional analysis for Germany. *Annals of Regional Science*. Vol. 45, pp. 55-85.
- Audretsch, D.B., Fritsch, M., 1994. On the measurement of entry rates. *Empirica*. Vol. 21, pp. 105-113.
- Audretsch, D.B., Fritsch, M., 2002. Growth regimes over time and space. *Regional Studies*. Vol. 36 (2), pp. 113-124.
- Audretsch, D., Keilbach, M., 2003. Entrepreneurship capital and economic performance. Discussion Paper 3678, Centre for Economic Policy Research.
- Bergmann, H., 2004. Gründungsaktivitäten im regionalen Kontext: Gründer, Gründungseinstellungen und Rahmenbedingungen in zehn deutschen Regionen. In: *Kölner Forschungen zur Wirtschafts- und Sozialgeographie*. Bd. 57. Köln.
- Boente, W., Heblich, S., Jarosch, M., 2008. Entrepreneurship capital, Knowledge spillovers and regional productivity: Some empirical evidence from European regions. Working Paper IAREG WP3/05. November.
- Bosma, N., Schutjens, V., 2011. Understanding regional variation in entrepreneurial activity and entrepreneurial attitude in Europe. *Annals of Regional Science*. 47(3), pp. 711-742.
- Chepurensko A. (2011). Entrepreneurship and SME Policies in the 21st Century – The example of Russia. In: Smallbone D., and Welter F. (eds.). Handbook of Research on Entrepreneurship Policies in Central and Eastern Europe. Cheltenham: Edward Elgar, p. 190-209.
- Chepurensko A., Gabelko M. V., Obratsova O., 2011. Early-stage entrepreneurial activity: an explanatory model for cross-country comparisons. Working papers by NRU HSE. Series WP1. No. 04.
- Doing Business in Russia 2012, at [http://www.russianlawonline.com/law-firms-publications/doing\\_business\\_russia\\_2012.pdf](http://www.russianlawonline.com/law-firms-publications/doing_business_russia_2012.pdf)
- Fisher R.A., 1936. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*. Vol. 7(2), pp. 179–188.

- Foelster, S., 2000. Do entrepreneurs create jobs? *Small Business Economics*. Vol. 14 (2), pp. 137-148.
- Frankfort-Nachmias, Ch., Leon-Guerrero, A., 2014. *Social Statistics for a Diverse Society*. 7-th ed.
- Friedman, J. H., 1989. Regularized discriminant analysis. *Journal of the American Statistical Association*. Vol. 84(405), pp. 165–175.
- Fritsch, M., Storey, D.B., 2014. Entrepreneurship in a regional context: Historical roots, recent developments and future challenges. *Regional Studies*. Vol. 48 (6), pp. 939–954.
- GEM 2011 Global Report, by Bosma, N., Sander Wennekers, S., Amoros, J.E., 2012, at <http://www.gemconsortium.org/docs/download/2200>
- GEM 2012 Global Report, by Xavier, S. R., Kelley, D., Kew, J., Herrington, M., Vorderwülbecke, A., 2013, at <http://www.gemconsortium.org/docs/2645/gem-2012-global-report>
- Hessels, J., van Stel, A., 2011. Entrepreneurship, export orientation, and economic growth. *Small Business Economics*. Vol. 37 (2), pp. 255-268.
- Hessels, J., van Gelderen, M., Thurik, R., 2008. Entrepreneurial aspirations, motivations, and their drivers. *Small Business Economics*, Vol. 31(3), pp. 323-339.
- Johannisson, B., 2007. Enacting local economic development – theoretical and methodological challenges. *Journal of Enterprising Communities: People and Places in the Global Economy*. Vol. 1(1), pp. 7–26.
- Keeble, D., Walker, S., 1994. New firms, small firms and dead firms: Spatial patterns and determinants in the United Kingdom. *Regional Studies*. Vol. 28, pp. 1-27.
- Koellinger, P., 2008. Why are some entrepreneurs more innovative than others? *Small Business Economics*. Vol. 31(1), pp. 21-37.
- Lee, S.Y., Florida, R., Acs, Z.J., 2004. Creativity and entrepreneurship: a regional analysis of new firm formation. *Regional Studies*. Vol. 38, pp. 879-891.
- Lim, T., Loh, W., Shih, Y., 2000. A comparison of prediction accuracy, complexity and training time of thirty-three old and new classification algorithms. *Machine*. Vol. 40, pp. 203–228.
- Lückgen, I., Oberschachtsiek, D., Sternberg, R., Wagner, J., 2004. Nascent entrepreneurs in German regions: Evidence from the regional entrepreneurship monitor (REM). *IZA Discussion Papers*, 1394, Institute for the Study of Labor (IZA).
- Martinez, A.M. Kak, A.C., 2001. PCA versus LDA. *IEEE transactions on pattern analysis and machine intelligence*. Vol. 23 (2), pp. 228–233.

- McLachlan, G. J., 2004. Discriminant analysis and statistical pattern recognition. *Wiley Interscience*. ISBN 0471691151 MR 1190469
- Naudé, W., Gries, T., Wood, E., Meintjies, A., 2008. Regional determinants of entrepreneurial start-ups in a developing country. *Entrepreneurship and Regional Development*, Vol. 20(2), pp. 111-124.
- Obraztsova O., Popovskaya E., 2012. Microdata of population surveys in the context of the business demography statistics in Russia. *Voprosy Statistiki* [Statistical Issues]. No. 12 [Russian].
- Reynolds, P., Bosma, N., Autio, E., Hunt, S., De Bono, N., Servais, I., et al., 2005. Global entrepreneurship monitor: Data collection design and implementation 1998-2003. *Small Business Economics*, Vol. 24(3), pp. 205-231.
- Reynolds, P. D., Storey, D. J., Westhead, P., 1994. Cross national comparison of the variation in new firm formation rates: an editorial overview. *Regional Studies*. Vol. 28, pp. 343–346.
- Rosstat. General SME Census 2010. Moscow 2012 (Russian).
- Rosstat. Russian National Accounts in 2004-2011. Moscow 2012 (Russian).
- Rosstat. Population employment survey – 2010. Moscow 2011 (Russian).
- Valliere, D., Peterson, R., 2009. Entrepreneurship and economic growth: Evidence from emerging and developed countries. *Entrepreneurship & Regional Development*. Vol. 21 (5-6), pp. 459-480.
- van Stel, A., Carree, M., Thurik, R., 2005. The effect of entrepreneurial activity on national economic growth. *Small Business Economics*. Vol. 24(3), pp. 311-321.
- Wagner, J., Sternberg, R., 2002. The role of the regional milieu for the decision to start a new firm: Empirical evidence for Germany. *IZA Discussion Papers*. 494, Institute for the Study of Labor (IZA).
- Wennekers, S., van Stel, A., Thurik, R., Reynolds, P., 2005. Nascent entrepreneurship and the level of economic development. *Small Business Economics*. Vol. 24(3), pp. 293-309.
- Wong, P., Ho, Y. & Autio, E., 2005. Entrepreneurship, innovation and economic growth: Evidence from GEM data. *Small Business Economics*. Vol. 24(3), pp. 335-350.

## APPENDICES

App. 1. Regional distribution of the share of opportunity-based early entrepreneurs (SOBE) level in Russia in 2011

Table A1.1 Main descriptive statistics of the SOBE level variation among regions

Distribution Indicators	
Mean, %	44,02
Mode,%	42,58
Median, %	39,22
The Standard Deviation, %	15,34
The Variation Coefficient	0,35
1-st Decile, %	26,33
9-th Decile, %	64,20
The Decile Differentiation	2,44
The 1-st Quartile, %	34,57
The 3-d Quartile, %	54,58
The ½ Quartile Variation	0,255
The Range of variation, %	75,00
The semi-quartile range	0,582
The share of regions where the SOBE level is less than Mode	0,443
The share of regions where the SOBE level is less than Mean <sub>Russia</sub>	0,468
The share of regions where the SOBE level is less than Mean <sub>GEM2011</sub>	0,190
The Lindberg's Excess	0,136
The Lindberg's Skewness	-0,260
The Pearson's Skewness	0,094

Table A1.2. Homogeneous groups of the regions by the SOBE level in 2011

Clusters in an Ordinary Scale	Space boundaries (Territories)	Region	SOBE level	Distance from a cluster centroid
Zero Level	Centre	Kostroma region	0,000	0,020
	Centre	Yaroslavl region	0,000	0,020
	North-West	Murmansk region	0,059	0,039
Below than Average Level	Centre	Tula region	0,167	0,018
	Centre	Tambov region	0,188	0,003
	Centre	Tver region	0,200	0,015
	South	Volgograd region	0,235	0,026
	Siberia	Altay Krai	0,250	0,011
	North-West	Novgorod region	0,278	0,016
	North	Dagestan Republic	0,282	0,021

	Caucasus			
Average Level	North Caucasus	Stavropol Region	0,333	0,064
	Volga region	Mordovia Republic	0,333	0,064
	North-West	Arkhangelsk region	0,350	0,048
	Centre	Ryazan region	0,353	0,045
	Centre	Kaluga region	0,379	0,018
	North-West	Komi Republic	0,381	0,017
	Volga region	Chuvash Republic	0,381	0,017
	South	Adygea Republic	0,385	0,013
	Volga region	Penza region	0,385	0,013
	Centre	Lipetsk region	0,389	0,009
	Centre	Orel region	0,389	0,009
	Ural	Sverdlovsk region	0,389	0,009
	Siberia	Irkutsk region	0,389	0,009
	Siberia	Khakassia Republic	0,391	0,006
	Siberia	Altai Republic	0,393	0,005
	Volga region	Saratov region	0,394	0,004
	Ural	Chelyabinsk region	0,395	0,003
	North-West	Pskov region	0,400	0,002
	North Caucasus	Kabardino-Balkaria Republic	0,400	0,002
	Volga region	Orenburg region	0,400	0,002
	Volga region	Perm Krai	0,400	0,002
	Ural	Kurgan region	0,400	0,002
	South	Kalmykia Republic	0,412	0,014
	Siberia	Tyva Republic	0,417	0,019
	Far East	Kamchatka Krai	0,417	0,019
	North-West	Karelia Republic	0,429	0,031
	North Caucasus	North Osetia - Alania Republic	0,429	0,031
	Far East	Sakhalin Region	0,435	0,037
	Centre	Belgorod region	0,444	0,047
	North-West	Saint Petersburg City	0,444	0,047
Siberia	Omsk region	0,444	0,047	
Far East	Magadan region	0,444	0,047	
Above Average Level	South	Astrakhan region	0,457	0,066
	Siberia	Trans-Baikal Territory	0,458	0,064
	North Caucasus	Karachaevo-Cherkessija Republic	0,462	0,061
	Siberia	Tomsk Oblast	0,462	0,061

	Ural	Tyumen region	0,464	0,059
	Centre	Bryansk region	0,469	0,054
	North-West	Kaliningrad region	0,471	0,052
	Siberia	Novosibirsk region	0,476	0,047
	Far East	Sakha (Yakutia) Republic	0,484	0,039
	Centre	Vladimir region	0,500	0,023
	Centre	Kursk region	0,500	0,023
	Volga region	Nizhny Novgorod region	0,500	0,023
	Far East	Amur Oblast	0,500	0,023
	Centre	Moscow region	0,514	0,009
	Ural	Khanty-Mansi Autonomous	0,519	0,004
	Volga region	Bashkortostan Republic	0,529	0,007
	Far East	Primorsky Krai	0,545	0,023
	Centre	Smolensk region	0,556	0,033
	Volga region	Udmurt Republic	0,556	0,033
	Siberia	Kemerovo region	0,567	0,044
	Volga region	Samara region	0,571	0,049
	Siberia	Krasnoyarsk Territory	0,579	0,056
	Centre	Ivanovo region	0,588	0,065
	North-West	Vologda region	0,591	0,068
High Level	Volga region	Kirov region	0,600	0,077
	Volga region	Ulyanovsk region	0,600	0,077
	Siberia	Buryatia Republic	0,600	0,077
	Centre	Voronezh region	0,625	0,066
	South	Krasnodar region	0,667	0,024
	Volga region	Mari El Republic	0,667	0,024
	Ural	Yamal-Nenets Autonomous	0,667	0,024
	South	Rostov region	0,682	0,009
	Centre	city of Moscow	0,692	0,001
	Far East	Jewish autonomous region	0,700	0,009
	Volga region	Tatarstan Republic	0,722	0,031
	Far East	Khabarovsk Krai	0,737	0,046
	North-West	Leningrad region	0,750	0,059

App. 2. List of indicators of the economic and social situation in Russian regions and their dynamics for 2008-2011 (Source: the Federal Statistical Service of the Russian Federation - Rosstat)

1	Population density, persons per sq.km in 2009
2	Actual final consumption of households per capita, thous. rubles in 2009
3	The percentage of urban population, % in 2009
4	Living space per capita (at the end of the year, square m) in 2009
5	The number of unemployed registered at public employment services, per one declared vacancy (at the end of the year) in 2009
6	The share of loss-making organizations (the total number of organizations) in 2009
7	Investments in fixed capital per capita, thous. rubles in 2009
8	Wear and tear of fixed capital in 2009
9	Demographic burden coefficient (per 1000 people of working age) in 2009
10	Infant mortality coefficient in 2009
11	Unemployment rate, in 2009
12	The economic activity rate (people employed or looking for work), % in 2009
13	Average income per capita, ths. rub in 2009
14	Average per capita money income (per month) to subsistence minimum ratio in 2009
15	The average annual number of employed in the economy in 2009
16	Number of recorded crimes per 10,000 people in 2009
17	Retail trade turnover, thous. rubles in 2009
18	The average size of the deposits of physical persons in RUR in Sberbank in 2009
19	The average size of the deposits of physical persons in foreign currency in Sberbank in 2009
20	The number of small businesses per 10,000 people in 2009
21	The number of people living below the poverty line in 2009
22	Average monthly nominal wages of workers, thous. rubles in 2009
23	Number of families, registered to need dwelling, % in 2009
24	Population density, persons per sq.km in 2010
25	Actual final consumption of households per capita, thous. rubles in 2010
26	The percentage of urban population, % in 2010
27	Living space per capita (at the end of the year, square m) in 2010
28	The number of unemployed people registered at public employment services, per one declared vacancy (at the end of the year, people) in 2010
29	The share of loss-making organizations (the total number of organizations) in 2010
30	Investments in fixed capital per capita, thous. rubles in 2010
31	Wear and tear of fixed capital in 2010
32	Demographic burden coefficient (per 1000 people of working age) in 2010
33	Infant mortality coefficient in 2010
34	Unemployment rate, % in 2010
35	The economic activity rate (people employed or looking for work), % in 2010
36	Average per capita money income (per month) , thous. rubles in 2010
37	Average per capita money income (per month) to subsistence minimum ratio in 2010
38	The average annual number of employed in the economy in 2010
39	Number of recorded crimes per 10,000 people in 2010
40	Retail trade turnover, thous. rubles in 2010

41	The average size of the deposits of physical persons in rubles in Sberbank in 2010
42	The average size of the deposits of physical persons in foreign currency in Sberbank in 2010
43	The number of small businesses per 10,000 people in 2010
44	The number of people living below the poverty line in 2010
45	Average monthly nominal wages of workers, thous. rubles in 2010
46	Number of families, registered to need dwelling, % in 2010
47	Population density, persons per sq.km in 2011
48	Actual final consumption of households per capita, thousand rubles in 2011
49	The percentage of urban population, % in 2011
50	Living space per capita (at the end of the year, square m) in 2011
51	The number of unemployed people registered at public employment services, per one declared vacancy (at the end of the year, people) in 2011
52	The share of loss-making organizations (the total number of organizations) in 2011
53	Investments in fixed capital per capita, thous. rubles in 2011
54	Wear and tear of fixed capital in 2011
55	Demographic burden coefficient (per 1000 people of working age) in 2011
56	Infant mortality coefficient in 2011
57	Unemployment rate, % in %_2011
58	The economic activity rate (people employed or looking for work), % in 2011
59	Average income per capita, ths. rubles in 2011
60	Average per capita money income (per month) to subsistence minimum ratio in 2011
61	The average annual number of employed in the economy in 2011
62	Number of recorded crimes per 10,000 people in 2011
63	Retail trade turnover, thous. rubles in 2011
64	The average size of the deposits of physical persons in RUR in Sberbank in 2011
65	The average size of the deposits of physical persons in foreign currency in Sberbank in 2011
66	The number of small businesses per 10,000 people in 2011
67	The number of people living below the poverty line in 2011
68	Average monthly nominal wages of workers, thous. rubles in 2011
69	Number of families, registered to need dwelling, %_2011
70	Population density, persons per sq.km, tempo in 2010
71	Actual final consumption of households per capita, tempo in 2010
72	The percentage of urban population, tempo in 2010
73	Living space per capita (at the end of the year, square m), tempo in 2010
74	The number of unemployed people registered at public employment services, per one declared vacancy (at the end of the year, people), tempo in 2010
75	The share of loss-making organizations (the total number of organizations), tempo in 2010
76	Investments in fixed capital per capita, tempo in 2010
77	Wear and tear of fixed capital, tempo in 2010
78	Demographic burden coefficient (per 1000 people of working age), tempo in 2010
79	Infant mortality coefficient, tempo in 2010
80	Unemployment rate, tempo in 2010
81	The economic activity rate (people employed or looking for work), tempo in 2010
82	Average income per capita, tempo in 2010
83	Average per capita money income (per month) to subsistence minimum ratio, tempo in 2010
84	The average annual number of employed in the economy, tempo in 2010
85	Number of recorded crimes per 10,000 people, tempo in 2010

86	Retail trade turnover, tempo in 2010
87	The average size of the deposits of physical persons in rubles in Sberbank, tempo in 2010
88	The average size of the deposits of physical persons in foreign currency in Sberbank, tempo in 2010
89	The number of small businesses per 10,000 people, tempo in 2010
90	The number of people living below the poverty line, tempo in 2010
91	Average monthly nominal wages of workers, thous. rubles, tempo in 2010
92	Number of families, registered to need dwelling _tempo_2010
93	Population density_tempo_2011
94	Actual final consumption of households per capita, tempo in 2011
95	The percentage of urban population, tempo in 2011
96	Living space per capita (at the end of the year, square m), tempo in 2011
97	The number of unemployed people registered at public employment services, per one declared vacancy (at the end of the year, people), tempo in 2011
98	The share of loss-making organizations (the total number of organizations), tempo in 2011
99	Investments in fixed capital per capita, tempo in 2011
100	Wear and tear of fixed capital, tempo in 2011
101	Demographic burden coefficient (per 1000 people of working age), tempo in 2011
102	Infant mortality coefficient, tempo in 2011
103	Unemployment rate, tempo in 2011
104	The economic activity rate (people employed or looking for work), tempo in 2011
105	Average income per capita, tempo in 2011
106	Average per capita money income (per month) to subsistence minimum ratio, tempo in 2011
107	The average annual number of employed in the economy, tempo in 2011
108	Number of recorded crimes per 10,000 people, tempo in 2011
109	Retail trade turnover, tempo in 2011
110	The average size of the deposits of physical persons in rubles in Sberbank, tempo in 2009
111	The average size of the deposits of physical persons in foreign currency in Sberbank, tempo in 2009
112	The number of small businesses per 10,000 people, tempo in 2011
113	The number of people living below the poverty line, tempo in 2011
114	Average monthly nominal wages of workers, thous. rubles, tempo in 2011
115	Number of families, registered to need dwelling _tempo_2011
116	Number of families, registered to need dwelling _tempo_2009
117	The average size of the deposits of physical persons in rubles in Sberbank, tempo in 2011
118	The average size of the deposits of physical persons in foreign currency in Sberbank, tempo in 2011
119	The share of households had an actual Internet access at home in 2010
120	The share of households had an actual Internet access, tempo in 2011
121	The share of households had a personal computer, tempo in 2011
122	The share of households had an actual Internet access at home in 2011
123	The share of households had a personal computer at home in 2011

*\*Source: Federal Statistical Service of the Russian Federation (Rosstat)*

App. 3. Structure matrix: discriminant loading of the discriminant functions\*\*

Discriminant Variables	Functions			
	1	2	3	4
The share of households had an actual Internet access at home in 2011	,673*	,069	,039	,010
The share of households had a personal computer at home in 2011	,487*	,063	,026	,008
Average monthly nominal wages of workers, thous. rubles, tempo in 2011 <sup>a</sup>	,223*	,124	,183	,096
Investments in fixed capital per capita, tempo in 2010 <sup>a</sup>	-,196*	-,073	-,020	-,030
Average monthly nominal wages of workers, thous. rubles, tempo in 2010 <sup>a</sup>	-,190*	-,052	-,108	,017
The number of small businesses per 10,000 people in 2010	,178*	-,085	,022	-,015
The average annual number of employed in the economy in 2009	-,156*	-,052	,070	-,022
The average annual number of employed in the economy in 2010 <sup>a</sup>	-,156*	-,052	,071	-,021
The average annual number of employed in the economy in 2011 <sup>a</sup>	-,153*	-,053	,072	-,019
Living space per capita (at the end of the year, square m)_2010	-,152*	-,021	,085	-,058
Living space per capita (at the end of the year, square m)_2011 <sup>a</sup>	-,152*	-,023	,084	-,059
The average size of the deposits of physical persons in foreign currency in Sberbank, tempo in 2010 <sup>a</sup>	,147*	,141	,038	,124
Unemployment rate, % in 2011 <sup>a</sup>	,139*	,026	,008	-,079
Number of recorded crimes per 10,000 people, tempo in 2010 <sup>a</sup>	,138*	,056	,120	-,077
Number of families, registered to need dwelling, % in 2009	,119*	-,091	-,027	-,023
Living space per capita (at the end of the year, square m)_2009	-,113*	,017	,025	-,085
Unemployment rate, % in 2010	,109*	-,015	-,047	-,022
Number of families, registered to need dwelling, % in 2010 <sup>a</sup>	,105*	-,055	-,021	-,021
Average per capita money income (per month) to subsistence minimum ratio_2009	-,105*	-,040	,085	,061
Unemployment rate, % in 2009	,100*	-,006	-,032	-,045
Population density, persons per sq.km in 2010	-,096*	-,094	,008	,080
Population density, persons per sq.km in 2011 <sup>a</sup>	-,096*	-,094	,008	,080
Population density, persons per sq.km in 2009 <sup>a</sup>	-,096*	-,094	,008	,081
Retail trade turnover, thous. rubles in 2011 <sup>a</sup>	-,089*	-,036	,073	,081
Number of families, registered to need dwelling, % in 2011 <sup>a</sup>	,087*	-,067	-,023	-,046
Infant mortality coefficient_tempo_2011 <sup>a</sup>	-,086*	,057	,002	-,034
Unemployment rate, tempo in 2010 <sup>a</sup>	,086*	-,024	-,062	-,052
Retail trade turnover, tempo in 2010 <sup>a</sup>	-,085*	-,039	-,068	-,036
Average per capita money income (per month) to subsistence minimum ratio_2010	-,081*	-,023	,070	,066
Wear and tear of fixed capital, tempo in 2011 <sup>a</sup>	-,065*	,064	,046	-,034
Average per capita money income (per month) to subsistence minimum ratio_tempo_2010 <sup>a</sup>	,046*	,039	,009	-,006
Number of families, registered to need dwelling, tempo in	-,044	,251*	,104	,050

2010 <sup>a</sup>				
The percentage of urban population, tempo in 2010 <sup>a</sup>	,005	,218*	-,031	,010
Demographic burden coefficient (per 1000 people of working age) in 2011 <sup>a</sup>	,021	,209*	-,179	,040
The average size of the deposits of physical persons in rubles in Sberbank, tempo in 2011 <sup>a</sup>	,165	,200*	-,064	-,068
The average size of the deposits of physical persons in rubles in Sberbank, tempo in 2010 <sup>a</sup>	,118	,197*	-,121	,041
Demographic burden coefficient (per 1000 people of working age) in 2010	,022	,189*	-,166	,061
Demographic burden coefficient (per 1000 people of working age) in 2009	,015	,178*	-,157	,079
The percentage of urban population, tempo in 2011 <sup>a</sup>	-,127	,175*	,004	-,074
Average monthly nominal wages of workers, thous. rubles in 2010	-,044	-,171*	,121	,029
Population density, tempo in 2011 <sup>a</sup>	-,011	-,170*	,052	,036
Average monthly nominal wages of workers, thous. rubles in 2011 <sup>a</sup>	-,034	-,165*	,131	,031
Number of families, registered to need dwelling, tempo in 2011 <sup>a</sup>	-,005	-,163*	,033	-,116
Average monthly nominal wages of workers, thous. Rub. in 2009	-,030	-,162*	,128	,024
The average annual number of employed in the economy, tempo in 2010 <sup>a</sup>	,025	-,158*	,106	,038
Average income per capita, thous. rubles in 2011 <sup>a</sup>	-,086	-,141*	,106	,063
Investments in fixed capital per capita, thous. rubles in 2010	-,108	-,141*	,069	,052
Average income per capita, thous. rubles in 2010	-,083	-,137*	,097	,061
Average income per capita, thous. rubles in 2009	-,073	-,132*	,101	,052
Investments in fixed capital per capita, thous. rubles in 2009	-,076	-,131*	,071	,044
Investments in fixed capital per capita, thous. rubles in 2011 <sup>a</sup>	-,101	-,129*	,069	,026
The average size of the deposits of physical persons in foreign currency in Sberbank, tempo in 2011 <sup>a</sup>	,062	,119*	-,088	-,050
Actual final consumption of households per capita, thousand rubles in 2011 <sup>a</sup>	-,002	-,112*	,025	,108
Actual final consumption of households per capita, thousand rubles in 2009	,011	-,111*	,024	,109
The number of people living below the poverty line, tempo in 2011 <sup>a</sup>	,010	,079*	-,017	-,071
Living space per capita (at the end of the year, square m)_tempo_2010 <sup>a</sup>	-,058	-,032	,274*	,009
The percentage of urban population, % in 2010 <sup>a</sup>	-,037	,010	,185*	-,025
The percentage of urban population, % in 2011 <sup>a</sup>	-,031	,002	,184*	-,022
The percentage of urban population, % in 2009	-,031	-,006	,183*	-,022
The economic activity rate (people employed or looking for work), % in 2009	,008	-,024	,179*	,032
Infant mortality coefficient_tempo_2010 <sup>a</sup>	-,060	,053	-,172*	-,097

The economic activity rate (people employed or looking for work), tempo in 2010 <sup>a</sup>	-,008	,006	-,166*	-,139
The share of loss-making organizations (the total number of organizations), tempo in 2011 <sup>a</sup>	,053	-,086	,161*	-,038
The number of small businesses per 10,000 people in 2010	,013	-,079	,159*	-,009
The number of small businesses per 10,000 people, tempo in 2011	-,110	-,058	,147*	-,051
The economic activity rate (people employed or looking for work), % in 2011 <sup>a</sup>	-,046	-,031	,142*	,088
Wear and tear of fixed capital in 2010	-,020	,094	-,141*	-,077
Wear and tear of fixed capital in 2011 <sup>a</sup>	-,050	,115	-,137*	-,090
Wear and tear of fixed capital in 2009	-,019	,075	-,131*	-,090
The economic activity rate (people employed or looking for work), % in 2010 <sup>a</sup>	,005	-,021	,130*	-,015
Retail trade turnover, tempo in 2011	-,003	-,033	-,126*	-,116
Average income per capita, tempo in 2010 <sup>a</sup>	-,086	,018	-,119*	,004
The number of small businesses per 10,000 people in 2009	,016	-,062	,118*	,064
The number of people living below the poverty line, tempo in 2010 <sup>a</sup>	,040	-,064	,110*	-,006
Living space per capita (at the end of the year, square m)_tempo_2011 <sup>a</sup>	-,026	-,051	-,105*	,053
Actual final consumption of households per capita, tempo in 2011	,095	,022	-,103*	,027
The number of small businesses per 10,000 people, tempo in 2010 <sup>a</sup>	,070	-,011	-,097*	,077
Number of recorded crimes per 10,000 people, tempo in 2011 <sup>a</sup>	,002	,051	,096*	,087
Actual final consumption of households per capita, tempo in 2010 <sup>a</sup>	,068	,014	-,072*	,051
Average income per capita, tempo in 2011	,026	-,004	,066*	,001
The average size of the deposits of physical persons in rubles in Sberbank in 2011	-,009	-,010	-,042*	-,018
The average size of the deposits of physical persons in foreign currency in Sberbank in 2011	-,004	-,025	,041*	,028
The number of unemployed people registered at public employment services, per one declared vacancy (at the end of the year, people) in 2011 <sup>a</sup>	,055	,097	,036	,277*
The share of loss-making organizations (the total number of organizations) in 2010	,025	-,029	-,048	,253*
The economic activity rate (people employed or looking for work), tempo in 2011 <sup>a</sup>	-,127	-,012	,008	,245*
The number of unemployed people registered at public employment services, per one declared vacancy (at the end of the year, people) in 2010	,045	,075	,030	,242*
The number of unemployed people registered at public employment services, per one declared vacancy (at the end of the year, people), tempo in 2011 <sup>a</sup>	-,151	,156	,041	,241*
Number of recorded crimes per 10,000 people in 2010	,042	-,035	,106	-,233*

The share of loss-making organizations (the total number of organizations) in 2011 <sup>a</sup>	,057	-,075	,062	,233*
Number of recorded crimes per 10,000 people in 2011	,036	-,022	,135	-,224*
The number of unemployed people registered at public employment services, per one declared vacancy (at the end of the year, people) in 2009	,068	,067	,008	,211*
Unemployment rate, tempo in 2011 <sup>a</sup>	,100	,133	,199	-,205*
Demographic burden coefficient (per 1000 people of working age), tempo in 2010 <sup>a</sup>	,062	,039	-,009	-,201*
Number of recorded crimes per 10,000 people in 2009	,013	-,052	,084	-,199*
The average annual number of employed in the economy, tempo in 2011 <sup>a</sup>	,080	-,033	,182	,192*
Investments in fixed capital per capita, tempo in 2011	,036	-,003	-,053	-,192*
Demographic burden coefficient (per 1000 people of working age), tempo in 2011 <sup>a</sup>	,013	,173	-,115	-,182*
Population density, tempo in 2010 <sup>a</sup>	,097	-,020	-,041	,176*
The share of loss-making organizations (the total number of organizations), tempo in 2010 <sup>a</sup>	-,019	-,085	-,142	,170*
Infant mortality coefficient_tempo_2009	,027	-,041	,053	,167*
The share of loss-making organizations (the total number of organizations) in 2009	,066	,038	,064	,151*
Infant mortality coefficient_2010	,007	-,030	-,004	,132*
The number of unemployed people registered at public employment services, per one declared vacancy (at the end of the year, people), tempo in 2010 <sup>a</sup>	,032	-,079	-,041	,129*
Actual final consumption of households per capita, thous. rubles in 2010	,000	-,115	,031	,118*
Average per capita money income (per month) to subsistence minimum ratio_2011 <sup>a</sup>	-,078	-,016	,002	,096*
Infant mortality coefficient in _2011 <sup>a</sup>	-,053	-,009	-,036	,096*
Retail trade turnover, thous. rubles in 2010	-,080	-,032	,086	,096*
Retail trade turnover, thous. rubles in 2009	-,063	-,018	,085	,092*
The average size of the deposits of physical persons in foreign currency in Sberbank in 2009	-,022	,029	-,067	-,086*
Wear and tear of fixed capital, tempo in 2010	,023	-,001	,017	,078*
The number of people living below the poverty line in 2009	,030	-,036	-,031	-,073*
The number of people living below the poverty line in 2011 <sup>a</sup>	,048	-,045	,001	-,072*
The average size of the deposits of physical persons in rubles in Sberbank in 2010	-,046	-,056	,035	,071*
The average size of the deposits of physical persons in rubles in Sberbank in 2009	-,051	-,061	,036	,067*
The number of people living below the poverty line in 2010	,044	-,057	,004	-,063*
Average per capita money income (per month) to subsistence minimum ratio_tempo_2011 <sup>a</sup>	,037	,033	-,025	,060*
The average size of the deposits of physical persons in foreign currency in Sberbank in 2010	,011	,010	-,007	-,012*

\* The maximum absolute value of the correlation between variable and the discriminant function.

\*\* The structure matrix represents the within-group correlations between the observed variables (the 123 continuous discriminating variables) and the dimensions created with the unobserved discriminant functions (dimensions, i.e. standardized canonical discriminant functions). When we use a stepwise method of variable selection, as we did in this problem, variables are entered in descending order of the absolute values of correlation inside the 1-st discriminant function.

a. This variable is not used in the analysis to support a conclusion that multicollinearity is not a problem for separating groups of regions. Multicollinearity is indicated by SPSS for step-wise FLDA by very small tolerance values for variables (because zero size of the tolerance means that the variable is a linear combination of other ones). The smallest tolerance for any variable included is 0.55 in present analysis.

---

### **Contact details and disclaimer**

**Alexander Chepureno**

Prof., Higher School of Economics, Moscow, [achepurenko@hse.ru](mailto:achepurenko@hse.ru)

**Any opinions or claims contained in this Working Paper do not necessarily reflect the views of HSE.**

**© Chepureno, Obraztsova, Elakhovsky, 2015**