

N.B.Davidson^{a)}, O.S.Mariev^{a)}

^{a)} Ural Federal University named after the first President of Russia B.N. Yeltsin

ESTIMATING THE IMPACT OF SPATIAL CONCENTRATION ON ENTERPRISE PERFORMANCE¹

The objective of this paper is to reveal the impact of spatial concentration of business in the Russian cities on enterprise productivity. The hypotheses are the following: urbanization level and home market potential positively affect enterprise performance; localization economies are positive and start decreasing after some point due to congestion and excessive competition; regional transport infrastructure, business climate and human capital positively affect enterprise performance. We use firm level data augmented with city and regional data. Fixed effects are applied in order to deal with endogeneity. Agglomeration economies are considered in the light of opportunities for knowledge spillovers, input sharing and labor market pooling. Our results confirm that agglomeration economies and home market potential are important for the enterprise performance. We find positive urbanization and diversity economies, while localization economies have an inverted U shape. Results can be used to improve regional policy. For instance, significance of home market potential emphasizes the importance of transport infrastructure. Significance of agglomeration effects implies that if a sufficiently large number of firms work in a city, performance of each firm improves.

Key words: agglomeration economies, localization, urbanization, diversity, home market potential, enterprise performance, Russia

1. Introduction

Regions and cities tend to differ in their initial opportunities and in level of economic development. Policy alternatives can favor either spatial concentration or equal geographical distribution of economic resources. The choice is affected by possible economic externalities from spatial concentration of economic activities. Therefore our *objective* is to shed some light on the impact of agglomeration levels and home market potential on enterprise performance. In order to account for various types of agglomeration economies, we study diversity, urbanization and localization economies. Home market potential, human capital, road infrastructure and business climate are taken into consideration. A number of agglomeration indices are analyzed, and a modified diversity index is suggested.

¹ The authors would like to thank the Russian Science Foundation for its support of research project №15-18-10014 «Projection of optimal socio-economic systems in turbulence of external and internal environment». We also express our gratitude to Hubert Jayet, Sergey Kadochnikov, Volodimir Vakhitov, Pavel Vorobyev, Nadezhda Kislyak for their valuable comments and ideas for further research. Any remaining errors are our own.

The determinants of enterprise productivity considered in this paper, particularly agglomeration economies and human capital are associated with education level. For example, education level affects *matching* on the labor market, i.e. firms' possibilities to find workers with needed qualifications, and workers' possibilities to find suitable jobs. Besides, education level is associated with *knowledge spillovers* among enterprises. These are two out of three *microeconomic foundations* of agglomeration economies. Next section of our paper discusses them in more detail. Another explanatory variable associated with educational level in our model is wage, which is used as proxy for human capital. Human capital is assumed to be affected by quality of education. Overall, our model captures the impact of education on enterprise productivity through two channels: agglomeration level and the level of human capital.

Econometric model elaborated in this research can be used to construct optimal spatial allocation of firms and to develop cluster policy. The next section is devoted to literature review. In section 3 data and methodology are discussed. Section 4 is devoted to the results, and the conclusions follow.

2. Literature review

Theoretical background. Spatial economics is a powerful branch of economic theory. It dates back to Weber (1929), and nowadays it attracts interest as New Economic Geography, starting with the works by Krugman, Fujita and Venables [1,2,3,4]. The research focus has been changing to reflect new economic realities. In the past researchers were occupied mainly with the optimal spatial allocation of resources that would minimize transportation costs. Later interest moved to the phenomenon of *agglomerations* associated with large cities or industrial clusters attracting human, natural, and financial resources. New economic geography explains agglomeration combining trade costs with scale economies [2].

Agglomeration is one of the central concepts in our research; in a broad sense it can be defined as concentration of economic activity in certain territories [5]. For example, Krugman describes spatial organization of economic activity in the United States, where the majority of population resides in the East Coast, although climate there is not the most favorable. He also mentions the pattern of economic activity location in Europe, where nighttime satellite photos indicate concentration in or near Belgium [2]. Marshall analyzed the supply side of agglomeration economies [5]. Particularly, he assumed that among the reasons for concentration of economic activity were 'the mysteries of trade' that became accessible to people in places of industrial concentration.

Marshall's line of thinking shows why the concept of agglomeration is important. Indeed, spatial concentration of economic activity associated with 'the mysteries of trade' leads to *agglomeration externalities or agglomeration economies*. They can be understood as positive and negative effects of scale and

scope resulting from economic activity concentration [6]. Agglomeration externalities play an important role in geographic allocation of resources because firms are likely to take them into account while choosing their location. Another concept used in our research is *home market potential*. This concept belongs to the New economic geography (NEG) models. According to the concept of home market potential, if two-region economy and the industries with imperfect competition are considered, then more than proportionate share of enterprises choose to locate in a region where local demand is larger [7].

Overall, enterprise performance is determined by various factors. Namely, there are factors inherent to place, such as geographical location and natural resources that can be classified as factors of the first nature. Besides, there are factors created by people, long term capital investments such as those into transport infrastructure, i.e. factors of the second nature. Finally, there are *agglomeration externalities*, i.e. factors of the third nature, which are the main focus of this research [8].

Rosenthal and Strange determine 3 scopes of agglomeration economies: industrial, geographical and temporal ones. In this research we focus on the industrial scope [9]. It reflects the industry boundaries of agglomeration economies: whether agglomeration economies are localized in one industry or they arise when different industries interact. In this respect, two broad types of agglomeration economies are known in the literature: localization and urbanization (diversity) economies.

Localization economies (MAR-externalities, firstly developed by Marshall and rediscovered by Arrow and Romer) are industry-specific externalities. They are associated with specialization, i.e. high concentration of economic activity in the same industry in a city. *Diversity economies* (Jacobs' externalities) are industry-universal externalities. They imply 'cross-fertilization' of different industries in a city and were explicitly formulated by Jane Jacobs [10]. They are externalities from diversity of economic activity in a city outside the own industry. *Urbanization economies* are associated with the city size, in other words, the number of people or the volume of economic activity in a city [9]. Diversity and urbanization economies are closely interrelated.

Among *diversity (urbanization) economies* associated with production and consumption are scale economies, common input market, decreasing transaction costs and additional opportunities, such as lower probability of unemployment and highly diversified goods and services [11]. Jacobs emphasized the importance of industrial diversity for social security and urban economic growth [12]. She claimed that small firms benefit more than the large ones from urban diversity, due to their stronger dependence on the market situation. Jacobs has also demonstrated that urban economic diversity is a critical condition for innovation [10].

Concerning economic foundations behind agglomeration processes, agglomeration economies are based on 3 principal mechanisms, or micro-

economic foundations: *sharing, matching and learning* [13]. Sharing mechanisms imply that agglomerations facilitate sharing of some common indivisible resources by firms, including infrastructure, variety of intermediate inputs, specialized labor, and risks. Matching mechanisms account for creation of pools of specialized workers, so the costs associated with training of workers decrease. Finally, learning mechanisms refer to the possibility of more intensive innovations in diverse environment of agglomerations stimulated by information spillovers [13].

Sharing externalities are in action both for localization and diversity measures. Matching mechanisms of agglomeration economies are reflected by the majority of localization indices, and to a less extent by diversity indices. Learning mechanisms are reflected more precisely by diversity indices than by localization ones. Overall, localization economies are explained by all three micro-foundations: sharing, matching and learning. Diversity economies are based primarily on learning mechanism. They are associated with demand as well as supply and explain advantages of life in a large city with its variety of goods and services [14].

As for the evidence about learning externalities, Audretsch concludes that location of firms in a close proximity remains important in spite of development of telecommunication technologies, partly due to tacit knowledge [15]. Dissemination of such type of knowledge takes place during face to face communication, which is more intensive within agglomerations [16]. These effects are particularly evident in the university cities, for example, in USA, where they became centers of industrial development for such branches as semiconductors, biotechnologies, software. Tacit knowledge is important in the contemporary times of communication technologies, as these technologies are complementary with face to face communication [17]. Another aspect of agglomeration effects are that people in large cities are more specialized, and therefore more productive, according to the logic of economic specialization suggested by Adam Smith (intensive margin). Besides, people are more numerous in the large cities, therefore they perform more tasks, becoming better as a team (extensive margin) [17].

Concerning the impact of agglomeration levels on productivity and wages, research on urban development suggests that there is a positive correlation between city size and wages, after controlling for ability. Explanation behind this is threefold. Firstly, there are high living costs in the cities, so wages are compensation. Secondly, highly qualified workforce is attracted to the cities, as a result of sorting. Thirdly, there are agglomeration externalities generated by interaction between the qualified employees and between the firms. As a result, productivity grows. Under imperfect competition wages can be defined as productivity minus profit margin of firms, and assuming that workers bear training costs to acquire the type of qualification

required by the firm, minus training costs. Therefore, as productivity grows with city size, wages can grow as well [17].

Along with positive agglomeration effects, high concentration of economic activities, for example, in large cities, may lead to crowding, high transport costs, ecological and social problems. There is a hypothesis of an inverted U-shape of agglomeration economies, proposed by Mills and Mirrlees [18, 19]. Mills develops a general equilibrium model with imperfect competition (monopoly in the goods market) and increasing returns to scale. He points out a possibility of congestion in the city beginning with a certain population density due to transportation costs and costs of housing. It implies diminishing agglomeration economies after a certain level of concentration, in other words, it reflects the hypothesis of inverted U-shaped agglomeration economies [18]. In our work, the inverted U shape of localization economies is confirmed.

Review of empirical results. In the overview of empirical works Beaudry and Schiffauerova document that researchers have found only positive MAR-externalities in 51 cases, both positive and negative ones in 11 cases, and in 20 cases they were non-significant [20]. Only positive Jacobs' externalities were found in 56 cases and non-significant in 46 cases; both positive and negative effects were found in 13 cases, and only negative effects in 9 cases.

De Groot et al. analyzed 31 papers and also found divergence in empirical results: both types of agglomeration externalities are positive in about the same number of cases as they are negative. For instance, they found that about half of the authors reveal negative Jacobs' externalities [21]. De Groot et al. have performed meta-analysis to study the determinants of the signs and significance levels of parameter estimates and came to the conclusion that the choice of the dependent variable, the control variables, and the construction of agglomeration indices affect their signs and significance levels [21]. Differences between countries and time periods also affect the parameter estimates. All this makes estimation of agglomeration effects context dependent [6].

Here is some evidence of positive agglomeration effects. Productivity premium received due to doubling of a city size, expressed through the population of a Standard Metropolitan Statistical Area (SMSA) for USA varies from 3 to 8% [9]. According to Sveikauskas, for an average industry, doubling of city size leads to productivity increase of 5.8% [22]. In the study by Okubo and Tomiura, productivity premium from being in a core region versus periphery was found to be 20–50% [23]. For Russia, based on data for 2005–2006, it was found that doubling city size increases firms' productivity by 5% [33]. Firm located within an agglomeration vs. outside an agglomeration has an increase in productivity by 46% [34].

Based on the US data for the year 1988, Ciccone and Hall found that doubling of employment density in a county leads to 6% increase of average labor productivity, where density is measured as the amount of labor per square foot; the results are aggregated from the county to state level [24]. Based on data

for Great Britain, the years 1998–2003, Anastassova found that estimated elasticity of hourly earnings with respect to employment density (the average number of full-time employees per square kilometer in a given district) is around 4 percent, both for district and county levels [25].

Applying plant-location fixed effects along with industry-time fixed effects, Henderson finds that localization externalities exist in high-tech, but not in machinery industries, and that diversity economies exist in machinery industry – corporate sector [26]. Martin et al. based on French data on enterprises for the years 1996–2004 find benefits from localization economies, but no benefits from diversity economies or competition effects [27]. They find the returns of localization economies to total factor productivity (TFP) to be 5–10%. They also find that firms internalize benefits from clustering, and therefore important effects from cluster policy should not be expected in developed countries. The dependent variable in their research is value added; explanatory variables are production factors, localization, diversity, urbanization and competition coefficients (in logarithms); they use time differencing approach. Time, employment area, industry and firm fixed effects are applied. While Henderson finds that it is the number of firms but not employees that generates externalities, Martin et al. come to the opposite conclusion [26, 27].

Henderson et al. find evidence of Marshall-Arrow-Romer (MAR) externalities in mature industries and evidence of both MAR and Jacobs' externalities in newer industries [28]. These conclusions are in line with concepts of agglomeration and product cycles. At the initial stage of development an industry benefits from a diverse environment within a metropolitan area. At a later stage, when the industry matures it benefits from location in smaller specialized cities. Localization economies dominate in such industries as textile, apparel, transport equipment, primary metals, food processing, pulp and paper. Diversity (urbanization) economies dominate in such industries as high-fashion apparel, upper-end publishing, and business services; they benefit from location in very large metropolitan areas.

At the enterprise level, diversity effects are usually found to be significant for new enterprises, while localization externalities are important for mature enterprises with standardized manufacturing production. Overall, 'the nursery cities' model implies that diversified cities and specialized cities co-exist, but are associated with different stages of industrial development [29]. Within the product cycle concept, it is assumed that firms move from an exploration (or search) stage to exploitation (or mass production) stage. At the first stage of development, location in the diversified environment is favorable for the firms, while at the second stage firms are more efficient when located in the specialized cities [6].

Overall, there is some evidence of positive agglomeration externalities, although the results are contradictory. In our research we distinguish between types of enterprises, and introduce enterprise fixed effects to enhance accuracy

of the empirical model. However, the issue of context dependence cannot be avoided completely. The next section is devoted to data and methodology used in this research.

3. Data description and methodology

Data. Research is based on the firm level data augmented with regional and city level data, for the years 2002–2008. The source of firm level data is SPARK-Interfax database, containing data on the organizational form, property, the year of foundation, location, revenue, labor, cost price and profit. The period under consideration belongs to the advanced stage of transition, after the 1997 new Privatization Law² and the 1998 financial crisis. The starting point of transition was the year 1991, the year of dissolution of the Soviet Union and turning from the centrally planned system to the market economy³.

The sample contains 7111 firms belonging to the manufacturing tradable industries⁴. The city level data is taken from Rosstat database on cities with population exceeding 100 thousand people. Regional data is collected by Rosstat too. Regional business climate indicators were constructed by the Analytical agency ‘Expert’⁵.

As there is a tendency for businesses that work in various regions of Russia to register in Moscow, the indicators of agglomeration in Moscow are probably exaggerated. Therefore researchers sometimes exclude Moscow (as well as St. Petersburg) from estimation. However, we did not exclude companies based in Moscow from our model, as the results with and without them proved not to differ significantly.

Measuring agglomeration. An important question in empirical studies is a choice of an appropriate measure for agglomeration level. There is a certain mechanism behind each indicator; for example, the number of plants included into analysis reflects possibility of knowledge circulation and other spillovers

² Database of legislation documents ‘Consultant Plus’

<http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=LAW;n=13731> (1991)

<http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=LAW;n=28103> (1997)

³ The first wave of transition in 1991-1995 is associated with price liberalization, mass privatization, financial system reforms and liberalization of external economic activities (Foundations of theory of transition economy, 1996).

⁴ SPARK: OKVED 10 (CA) – 37(DN), 40(E), 41(E) – i.e. whole E, 45(whole F), 60(part of I), 65(part of J), K and ‘high tech’ industries. There are 15609 firms belonging to the manufacturing tradable industries in the database, but only for 7111 of these firms city level data is present.

⁵ Analytical agency ‘Expert’ <http://www.raexpert.ru/ratings/regions/ratingclass/>. The results of Ranking of investment attractiveness of the Russian regions are published in the journal ‘Expert’ annually since 1996.

between firms [27]. Below we discuss indicators of agglomeration: *localization*, *urbanization* and *diversity* coefficients.

To analyze *localization*, the concepts of the *level of economic activity* in a city and *specialization* of a city can be used. These concepts reflect different aspects of localization, i.e. indices showing the level of economic activity reflect possibility of close interaction between firms in agglomerations. Meanwhile, specialization indices allow comparing development of various industries in different geographical locations and can be used, for example, to explain trade flows between them. In this research, the level of economic activity is considered, in order to capture agglomeration economies arising from interaction between firms. We use the following *localization coefficient* that reflects the *level of economic activity*:

$$\ln(\text{loc})_{it}^{jz} = \ln(\text{revenue}_i^{jz} - \text{revenue}_{it}^{jz} + 1), \quad (1)$$

where revenue_i^{jz} – the revenue of all firms belonging to an industry j and located in a city z ; revenue_{it}^{jz} – revenue of a firm i belonging to an industry j and located in a city z ; t is time [27, 30]. Localization can be measured based on various industry aggregation levels, while 3-digit aggregation level is basic in the literature [20]. In this research we also use 3-digit aggregation level.

As it was mentioned in the literature review, diversity and urbanization economies are closely interrelated. They are associated with total urban area population, total local employment, the number of industries in a city etc. Diversity and urbanization levels can be measured using the following indices.

Urbanization coefficient measures total revenue of firms belonging to all industries in a city, except for the industry under consideration:

$$\ln(\text{urb})_t^z = \ln(\text{revenue}_t^z - \text{revenue}_t^{jz} + 1) \quad (2)$$

where revenue_t^{jz} – revenue of all firms belonging to an industry j and located in a city z ; revenue_t^z – revenue of all firms in a city z ; t is time.

Diversity coefficient. Among the indices that measure diversity are Herfindahl–Hirschman index, Gini index of diversity, Theil index, Ellison–Glaeser index, and the Duranton and Overman index (continuous approach) [31, 26, 27]. A typical problem for diversity indices consists in their instability to elimination of the dominant industry, i.e. they tend to reflect inequality rather than variety of industries. In order to deal with this problem, we analyze diversity along two dimensions. The first dimension, *variety*, shows how many different industries exist in a city. The second one, *inequality*, reflects how evenly they are distributed [30]. The following diversity coefficient is used in this research:

$$div_i^z = \frac{\sum_{j=1}^s \left(\frac{revenue_t^{jz}}{revenue_t^z} \right)^{\frac{1}{s}} - 1}{s^{\frac{1}{s}} - 1}. \quad (3)$$

This coefficient accounts both for variety and for inequality [30]. It takes on the value 1 if the distribution of revenue among industries is perfectly equal, and the value 0 if all revenue is concentrated in one industry. In this paper, we suggest an index where s is the number of industries in a country. Within this approach, variety is reflected by comparing industrial structure in the city and in the country.

The indicators of agglomeration levels discussed above show that there is positive correlation between diversity and city size (the total revenue of enterprises in a city). Besides, in larger cities there is not only greater diversity but also greater localization levels, i.e. more enterprises of the same industry.

To summarize, *localization, urbanization and diversity economies* draw our attention to the processes taking place on supply side. We are also interested in demand side that an enterprise faces. Therefore *Home market potential* is introduced into our analysis. To construct the variable *Home market potential* (HMP) of a city, we consider potential demand in a city itself and in all other cities in the country, weighed by distances. Potential demand *in a city* is measured with total revenue of firms in a city.

Estimation of the agglomeration economies. Measuring agglomeration levels with urbanization, diversity and localization indices discussed above is one of the steps in constructing an econometric model to estimate the agglomeration economies. The next step is finding an appropriate measure of economic performance. And finally, the impact of agglomeration levels on economic performance is estimated.

Performance measures. Agglomeration economies are measured as productivity gains on a firm level. A specification based on Cobb-Douglas production function is used here, with total factor productivity of a firm (TFP) affected by agglomeration levels [9], home market potential of a city, human capital level, transport infrastructure, and business environment. Cobb-Douglas function proved to perform better than Translog function. We are interested in the factors enhancing productivity of enterprises, i.e. factors that allow an enterprise to produce more, using the same amount of labor and capital. Dependent variable is logarithm of enterprise revenue [26, 32].

Econometric model. The following econometric model is estimated:

$$\begin{aligned} \ln(revenue)_{ii}^{jz} = & \beta_0 + \beta_1 \ln(capital)_{ii}^{jz} + \beta_2 \ln(labour)_{ii}^{jz} + \beta_3 core_t^{jz} + \beta_4 [core_t^{jz}]^2 + \\ & + \beta_5 \ln(loc)_{ii}^{jz} + \beta_6 [\ln(loc)_{ii}^{jz}]^2 + \beta_7 \ln(urb)_t^{jz} + \beta_8 \ln(hmp)_t^z + \beta_9 \ln(wage)_t^z + \\ & + \beta_{10} \ln(a_road)_t^r + \beta_{11} \ln(rw_road)_t^r + \beta_{12} \ln(busnenvrisk)_t^r + \varphi_i + \varepsilon_{ii}, \end{aligned} \quad (4)$$

where j is industry index, z is city index, r is regional index, and i is firm index.

Estimation is done by regressing the logarithm of enterprise revenue on the enterprise characteristics, as well as on market size, agglomeration indices, and several other characteristics of territories. Explanatory variables used in equation (4) are presented in Table 1 below. Keeping in mind large size of the Russian regions, agglomeration indices and Home market potential were calculated on a city level. All agglomeration coefficients ('core', localization index, diversity and urbanization indices) are calculated based on 3-digit level of OKVED classification.

Robust standard errors were used to deal with the heteroscedasticity problem, as normality and homoscedasticity of the random error terms were not assumed. Based on panel data, specific features of industries and territories are reflected in the enterprise fixed effects, as in the sample firms do not change location or industry. Fixed effects are applied to deal with endogeneity. Hausman specification test showed that the fixed effects model is preferable over random effects method of panel data estimation. The models with industry level, regional level, and time fixed effects were also tested. Enterprise fixed effects were found to be the most relevant here.

Table 1.

Variables used in the econometric analysis

Variable	Definition
Enterprise level characteristics.	
$\ln(revenue)_{it}^{jz}$	Logarithm of enterprise revenue (revenue is in roubles).
$\ln(capital)_{it}^{jz}$	Logarithm of fixed assets (fixed assets are in roubles).
$\ln(labour)_{it}^{jz}$	Logarithm of labor force (number of employees).
Agglomeration indices. City level.	
$\ln(loc)_{it}^{jz}$	Logarithm of localization coefficient.
$core_t^{jz}$	A share of an industry j in the total revenue in a city z .
$\ln(div)_t^z$	Logarithm of diversity coefficient.
$\ln(urb)_t^{jz}$	Logarithm of urbanization coefficient.
Home market potential. City level: $HMP_{city} = \ln(div)_t^z + \ln(hmpcity)_t^z$.	
$\ln(hmpcity)_t^z$	$hmpcity_t^z$ is total revenue in all industries in all cities other than the city where the firm is located, divided into the distances between the

Variable	Definition
	city Z and the other cities: $hmpcity_t^z = \sum_{z \neq Z} \frac{revenue_t^z}{dist_{z,z}}$, where $dist_{z,z}$ are the physical distances between the cities; logarithm of this variable is used.
City characteristics.	
$\ln(wagecity)_t^z$	Logarithm of average monthly nominal wage, payroll, roubles (proxy for human capital).
Regional transport infrastructure.	
$\ln(a_road)_t^r$	Logarithm of automobile road density in region r , end of year, km of roads per 1000 sq km of territory.
$\ln(rw_road)_t^r$	Logarithm of railroad density in region r , end of year, km of roads per 1000 sq km of territory.
Regional investment climate (Indices by analytical agency ‘Expert’) ⁶ .	
$\ln(busnenvrisk)_t^r$	Logarithm of business environment risks, region r , time t (Regional investment risk by the Analytical agency Expert).

Localization loc_{jz} in a quadratic form is included into regression to test the inverted U shape of agglomeration economies suggested in the literature. This form seems to be more relevant for localization economies than for the urbanization ones, particularly due to competition among the firms of the same industry for labor force and for the other resources.

To measure diversity, the index presented in equation (3) above is applied. Explanatory variables $\ln(div)_t^z$ and $\ln(urb)_t^{jz}$ are used interchangeably in the regression, as cities with higher urbanization levels tend to be more diversified; correlation between these coefficients is about 90%.

To analyze the issues connected with the dominating role of an industry in a city we use a variable $core_{jz} = \frac{revenue_{jz}}{revenue_z}$, a share of an industry j in the total revenue in a city z .

The variable accounting for regional business environment risks $\ln(busnenvrisk)$ is included into the model, based on the assumption that the higher are the risks in the region, the more time entrepreneurs and managers spend to deal with these risks, and the less time is devoted to the production process. Moreover, higher risks divert attention from all kinds of innovations, such as improvement of products, technologies, and management processes.

⁶ Analytical agency ‘Expert’ <http://www.raexpert.ru/ratings/regions/ratingclass/>.

The results received using the approach discussed above, are presented in the next section.

4. Estimation results

Equation (4) was estimated using panel data model with enterprise fixed effects. The results of econometric testing of the equation (4) are presented in Table 2 below. The model was estimated for all firms belonging to the industries that produce tradable goods ('all tradables' – models 1 and 2), for the 'old' firms (models 3 and 4), and for the 'young' firms (models 5 and 6)⁷. For each type of firms the model was estimated with urbanization coefficient (models 1, 3 and 5) and with diversity coefficient (models 2, 4 and 6) interchangeably.

Table 2.

The impact of spatial concentration on enterprise performance
Dependent variable: $\ln(\text{revenue})$

	(1) All tradables-1	(2) All tradables-2	(3) 'Old' firms-1	(4) 'Old' firms-2	(5) 'Young' firms-1	(6) 'Young' firms-2
$\ln(\text{capital})$	0.234*** (22.59)	0.236*** (22.68)	0.233*** (15.81)	0.234*** (15.86)	0.236*** (16.16)	0.237*** (16.23)
$\ln(\text{labor})$	0.428*** (32.00)	0.431*** (32.07)	0.437*** (24.32)	0.439*** (24.29)	0.417*** (20.82)	0.421*** (20.98)
core	9.574*** (15.31)	8.738*** (14.30)	9.132*** (12.43)	8.305*** (11.65)	10.24*** (9.47)	9.428*** (8.81)
core ²	-6.628*** (-8.20)	-6.258*** (-7.79)	-6.612*** (-5.94)	-6.152*** (-5.59)	-6.787*** (-5.52)	-6.577*** (-5.35)
$\ln(\text{loc})$	0.0312*** (4.42)	0.0287*** (4.01)	0.0254*** (3.12)	0.0223*** (2.71)	0.0399*** (3.07)	0.0389*** (2.95)
$(\ln(\text{loc}))^2$	-0.00126*** (-3.37)	-0.00106*** (-2.77)	-0.000874** (-2.07)	-0.000651 (-1.52)	- (-2.69)	-0.00175** (-2.41)
$\ln(\text{urb})$	0.282*** (10.69)		0.244*** (7.61)		0.338*** (7.51)	
$\ln(\text{div})$		0.565*** (5.15)		0.424*** (3.39)		0.824*** (3.90)
$\ln(\text{hmcacity})$	0.451***	0.585***	0.440***	0.534***	0.461***	0.667***

⁷ 'Old enterprises' are enterprises created during planned economy. 'Young enterprises' are those founded starting with the year 1995, the end of the check privatization.

	(1) All tradables-1	(2) All tradables-2	(3) 'Old' firms-1	(4) 'Old' firms-2	(5) 'Young' firms-1	(6) 'Young' firms-2
	(8.44)	(11.53)	(6.23)	(7.89)	(5.68)	(8.92)
ln(wagecity)	-0.000303 (-0.01)	0.118*** (2.86)	0.0346 (0.65)	0.151*** (2.76)	-0.0468 (-0.77)	0.0662 (1.07)
ln(a_road)	0.0445* (1.66)	0.0652** (2.39)	0.0463 (1.39)	0.0580* (1.72)	0.0384 (0.84)	0.0780* (1.66)
ln(rw_road)	-0.466 (-1.54)	-0.500 (-1.57)	-0.260 (-0.69)	-0.274 (-0.69)	-0.937* (-1.81)	-1.030* (-1.94)
ln(busnenvrisk)	-0.00805 (-1.53)	-0.0166*** (-3.12)	-0.0126** (-2.10)	-0.0193*** (-3.19)	0.000830 (0.08)	-0.0115 (-1.09)
_cons	-3.588* (-1.90)	-0.0152 (-0.01)	-3.882 (-1.61)	-0.511 (-0.20)	-2.163 (-0.70)	1.673 (0.53)
firm fixed effects	yes	yes	yes	yes	yes	yes
<i>N</i>	49779	49779	32753	32753	17026	17026
adj. <i>R</i> ²	0.428	0.425	0.424	0.421	0.433	0.430
sigma_u	1.153	1.172	1.080	1.087	1.382	1.423
sigma_e	0.522	0.523	0.484	0.485	0.589	0.590
rho	0.830	0.834	0.833	0.834	0.846	0.853

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

For years 2002–2008, i.e. 7 years overall, there are 49779 observations for the firms belonging to tradable industries, comprising 32753 ‘old’ firms and 17026 for the ‘young’ firms. For each year there are 7111, 4679, and 2432 observations respectively.

Adjusted R squared (adj. R^2) shows that the variables included into the model explain from 42.1% to 43.3% of variation in the enterprise revenue. Therefore, the model gives substantial information on the determinants of enterprise revenue. Rho, the fraction of variance due to the individual effect u_i , or in other words due to differences across panels, is 83% to 85.3% in the models presented in Table 2⁸.

The results presented in Table 2 above show that localization, urbanization, diversity and HMP are significant. A share of an industry j in the total revenue in a city z (‘core’ coefficient) and localization coefficient have an

⁸ Rho is calculated according to the following formula: $\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$, where σ_u or sigma_u is standard deviation of residuals within groups u_i ; σ_e or sigma_e is standard deviation of residuals (overall error term) e_i [35].

inverted U shape. 1% increase in 'core' leads to 0.087–0.096% increase in enterprise revenue in Models (1)–(2). However, when 'core' reaches $8.738/(2*6.258)=0.698$, further 1% increase in 'core' leads to decrease in revenue by 0.063% compared to the optimum level of 'core', i.e. the 'core' effect stays positive but decreases. For Models (1) the threshold is 0.72.

As for localization coefficient, its increase by 1% leads to increase in enterprise revenue by around 0.029–0.03%. In case of Model (2), when the level of localization $0.0287/(2*0.00106)=13.538$ is reached i.e. a very low level of total revenue in industry in city equal to $\exp(13.538)=757,467$ roubles, then each 1% further increase in localization starts decreasing enterprise revenue by 0.00065%. In other words, the effect decreases although stays positive until $0.0287/(0.00106)=27.075$, i.e. $\exp(27.075)=0.57$ bn roubles⁹. In Model (1) the optimum localization level is 12.381.

Economic activity in industries other than that where a firm works is reflected by diversity and urbanization levels interchangeably. The values of the coefficients are robust to the changes in specification. Diversity is significant and positive, showing that 1% increase in diversity leads to 0.565% increase in enterprise revenue. It means that if we compare a 'Center' with diversity approaching to 1 and a 'Periphery' with diversity approaching to 0, i.e. were diversity would be twice lower, the difference in productivity would approach 56.5%. Urbanization is significant and positive too, 1% increase in urbanization leading to 0.28% increase in enterprise revenue. The results are consistent with the existing literature discussed in section 2 above.

HMP on city level consists of total revenue of firms in a city, excluding the revenue of the firm itself (urbanization coefficient) and the total revenue of firms in the other cities, the distances between cities being taken into account. As for HMP arising from the other cities, its increase by 1% leads to increase in enterprise revenue by 0.451–0.585% in specifications (1)–(2). The effect of wage is positive in specification with diversity index, possibly implying that firms benefit from human capital quality associated with higher wages. The effect of the automobile road density is positive in the specifications where it is significant. In specification with diversity coefficient, Models (2) and (4), business environment risk index is significant on 99% confidence interval showing that 1% increase in the regional business environment risk leads to decrease in the firm's revenue by 0.017%. In Model (3) it is significant on 95% confidence interval and stays negative.

We also compare agglomeration economies for the 'old' and 'young' firms in Russia. Agglomeration effects and HMP proved to be present both for the 'young' firms and for the firms established before privatization, i.e. 'old' firms. However, these effects are stronger for the 'young' firms; they benefit relatively more not only from diversity and urbanization economies, but also

⁹ Billion according to the long scale.

from localization economies. Specific features of agglomeration effects for different types of firms found based on data for the other countries were discussed in section 2 above.

Estimation issues. Endogeneity problem arises because explanatory variables may correlate with an error term. In this case OLS-estimates are biased. Endogeneity may occur for two major reasons [27].

Unobserved heterogeneity: there are some unobserved factors which affect both the output of enterprise and some explanatory variables. Because of this, the error term is correlated with explanatory variables, and explanatory variables capture part of the effect of the unobserved factors, leading to potential bias in coefficients estimated using the OLS model. One of the ways to address unobserved heterogeneity using panel data is individual fixed effects [26]. As enterprises do not change location and industry over time, these fixed effects capture specific features of industries and locations as well.

Simultaneity. A possible source of simultaneity comes from an assumption that firms choose their location considering potential agglomeration advantages or disadvantages. Successful businesses choose the cities where productivity is higher, and are therefore disproportionately found in agglomeration areas. Another potential source of simultaneity is productivity shocks that affect labor and capital that enterprise uses, and the other enterprises in a city or region, i.e. localization and urbanization or diversity levels [27].

The sample under consideration includes firms created under central planning (the ‘old’ firms). For them simultaneity problem should not arise because their location was not likely to be motivated by efficiency. We compare distribution of ‘old’ and ‘young’ enterprises along localization levels in the cities. It is assumed that if the distribution is similar, then simultaneity problem should not arise for the ‘young’ enterprises as well. The results are presented in the Figure 1 below.

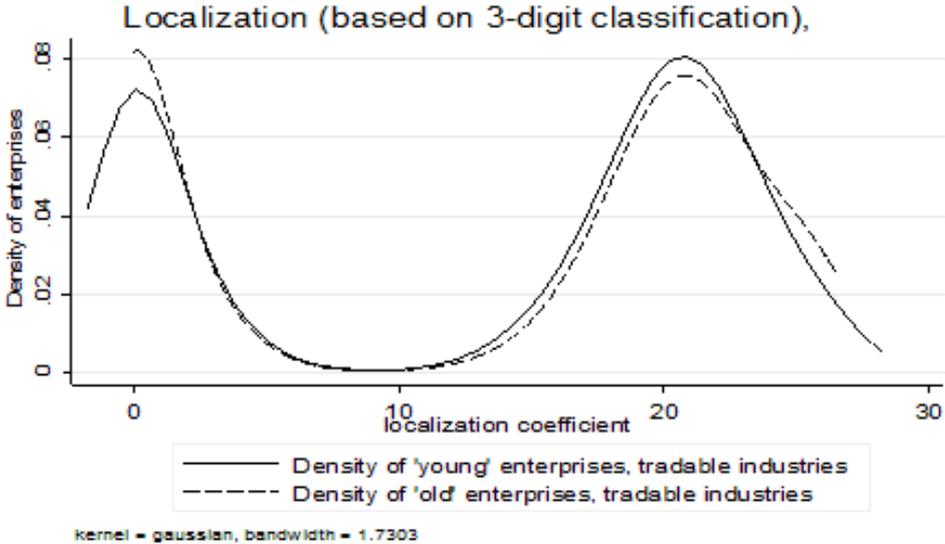


Figure 1. Actual distribution of 'old' and 'young' firms by localization level

Figure 1 above shows that distribution of 'old' and 'young' enterprises along localization levels is very similar. Therefore we suppose that simultaneity problem does not occur for the whole sample of firms. In order to check this assumption, we used lagged agglomeration variables as instruments for the agglomeration variables, and applied 2SLS and GMM procedures. The results remained robust.

5. Conclusion

In this paper we have analyzed the factors important for high productivity of enterprises. Among other factors, we studied the impact of agglomeration and home market potential (HMP) on the enterprise performance. Localization economies were found to have an inverted U shape for the majority of specifications. Urbanization economies, diversity economies and HMP proved to be positive and significant. As HMP is related to distance and transport costs, improvement of transport infrastructure and decrease in communications costs would increase positive effects associated with market access.

Taking a closer look at transport infrastructure, automobile road density is positive in the specifications where it is significant. Regional business risk measured by the Analytical agency 'Expert' has a negative sign in the specifications where it is significant. In other words, along with agglomeration economies and HMP, transport infrastructure and business climate are important for successful work of enterprises. To observe the effect of human capital on enterprise productivity, we have chosen wage as a proxy of human capital quality. The effect of wage on productivity is positive in the specifications where it is significant.

As the next step, it would be useful to analyze agglomeration economies and the impact of HMP for various industries, cities and firms, in order to have more precise conclusions for regional policy.

References

1. Weber, A. (1929), *The theory of the location of industries*, Chicago University Press, Chicago (English translation from the German version published in 1909), 255.
2. Krugman, P. (1991), Increasing returns and economic geography, *Journal of Political Economy*, 99, 483–499.
3. Fujita, M. (1988), A monopolistic competition model of spatial agglomeration: a differentiated product approach, *Regional Science and Urban Economics* 18, 87–124.

4. Venables, A.J. (1996), Equilibrium locations of vertically linked industries, *International Economic Review*, 37, 341–359.
5. Marshall, A. (1920), Principles of Economics, 8th edition, London, Macmillan and Co. Original edition 1890, 627.
6. Neffke, F. (2009), Productive places. The influence of technological change and relatedness on agglomeration externalities, Faculty of Geosciences, Utrecht University, the Netherlands, 199.
7. Ottaviano, G. and J. Thisse (2004), Agglomeration and Economic Geography, in V. Henderson and J. Thisse, eds. *Handbook of Urban and Regional Economics*, 4, 2563–2608.
8. Mikhailova, T. (2013), Looking for multiple equilibria in Russian urban system. Presentation at the 2nd International Conference ‘Industrial Organization and Spatial Economics’, St. Petersburg, Center for Market Studies and Spatial Economics National Research University ‘Higher School of Economics’, 27–30 June, 48.
9. Rosenthal, S. S. and W. C. Strange (2004), Evidence on the nature and sources of agglomeration economies, in V. Henderson and J. Thisse, eds. *Handbook of Urban and Regional Economics*, 4, 2119-2171.
10. Jacobs, J. (1969), The economies of cities, New York, Random House, 288.
11. Quigley (1998), Urban Diversity and Economic Growth, *Journal of Economic Perspectives*, 12(2) Spring, 127–138.
12. Jacobs, J. (1961), The death and life of great American cities, New York, Random House, 458.
13. Duranton, G. and D. Puga (2004), Micro-foundations of urban agglomeration economies. In V. Henderson and J. Thisse, eds. *Handbook of Urban and Regional Economics*, 4, 2063-2117.
14. Ottaviano, G. and J. Thisse (2004), Agglomeration and Economic Geography, in V. Henderson and J. Thisse, eds. *Handbook of Urban and Regional Economics*, 4, 2563-2608.
15. Audretsch, D. B. (1998), Agglomeration and the location of innovative activity, *CEPR Discussion Paper No. 1974*, 29.
16. Ellison, G. and E. L. Glaeser (1997), Geographic concentration in U.S. manufacturing industries: a dartboard approach, *The Journal of Political Economy*, 105(5), 889–927.
17. Robert-Nicoud, F. (2013), Agglomeration economies: matching and sharing, *Presentation at CMSSE Summer School in Spatial Economics and Imperfect Markets: empirical aspects. Pushkin*, 3–16 July, 54.
18. Mills, E. S. (1967), An aggregative model of resource allocation in a metropolitan area, *American Economic Review Papers and Proceedings*, 57(2), 197–210.
19. Mirrlees, J. A. (1972), The optimum city, *Swedish Journal of Economics*, 74(1), 114–135.

20. Beaudry, C. and A. Schifffauerova (2009), Who's right, Marshall or Jacobs? The localization versus urbanization debate, *Research Policy* 38, 318–337.
21. De Groot, H., Poot, J., and M.J. Smit (2009), Agglomeration externalities, innovation and regional growth: theoretical perspectives and meta-analysis. In: *Handbook of Regional Growth and Development Theories*, Capello R., and Nijkamp P. (eds.). Cheltenham: Edward Elgar, 256–281.
22. Sveikauskas L. (1975), The Productivity of Cities, *The Quarterly Journal of Economics*, 89(3), 393–413.
23. Okubo, T. and E. Tomiura (2010), Productivity distribution, firm heterogeneity, and agglomeration: evidence from firm-level data, *Discussion paper* No 10017, Research Institute of Economy, Trade and Industry (RIETI), 25.
24. Ciccone, A. and R. E. Hall (1996), Productivity and the density of economic activity, *American Economic Review*, 86(1), 54–70.
25. Anastassova, L. (2006), Productivity differences and agglomeration across districts in Great Britain. *CERGE-EI Working Paper* No. 289, 34.
26. Henderson, J.V. (2003), Marshall's scale economies, *Journal of Urban Economics*, 53, 1–28.
27. Martin, P., Mayer, T. and F. Mayneris (2011), Spatial concentration and plant-level productivity in France, *Journal of Urban Economics*, 69, 182–195.
28. Henderson, J.V., A. Kuncoro and M. Turner (1995), Industrial development in cities, *Journal of Political Economy* 103, 1067–1085.
29. Duranton, G. and D. Puga (2001), Nursery cities, *American Economic Review*, 91, 1454–1477.
30. Vorobyev P.V., Davidson N.B., Kislyak N.V., Kuznetsov P.D. (2014), Industrial diversity and concentration in the Russian cities as factors of economic efficiency, *Vestnik UrFU*, No 6, 4–18.
31. Combes, P.-P., Mayer, T. and J.-F. Thisse (2008), *Economic geography. The integration of regions and nations*, Princeton University Press, Princeton and Oxford, 399.
32. Vakhitov V. (2008), *Agglomeration economies and geographic concentration of manufacturing in Ukraine*, Dissertation, Director: Dr. Christopher Bollinger, Professor of Economics, Lexington, Kentucky, 158.
33. *Russian Manufacturing Revisited: two Rounds of Surveying Russian Enterprises in 2005–2009*, *HSE Policy Paper*, State University – Higher School of Economics, Moscow, 2010, 61.
34. Gonchar, K. (2010), *Urban agglomeration effects and company productivity in Russia: Empirical Evidence Based on Manufacturing Industry Survey*, *Report at ISNIE*, 13.
35. Baum, Ch. (2006), *An introduction to modern econometrics using Stata*, Texas, A Stata Press Publication, 363.

Davidson Natalia Borisovna (Yekaterinburg, Russia) — PhD in Economics, Senior Lecturer of the Chair of International Economics of Graduate School of Economics and Management, Ural Federal University named after the first President of Russia B.N. Yeltsin (19, Mira St., Yekaterinburg, 620002, Russia, e-mail: natalya.davidson@gmail.com).

Mariev Oleg Svyatoslavovitch (Yekaterinburg, Russia) — PhD in Economics, Head of the Chair of Econometrics and Statistics of Graduate School of Economics and Management, Ural Federal University named after the first President of Russia B.N. Yeltsin (19, Mira St., Yekaterinburg, 620002, Russia, e-mail: o.s.mariev@urfu.ru).