NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Alexander S. Skorobogatov

SPATIAL EQUILIBRIUM APPROACH TO THE ANALYSIS OF INCOME DIFFERENTIALS ACROSS RUSSIAN CITIES

BASIC RESEARCH PROGRAM
WORKING PAPERS

SERIES: ECONOMICS
WP BRP 149/EC/2016

Alexander S. Skorobogatov[*]

SPATIAL EQUILIBRIUM APPROACH TO THE ANALYSIS OF INCOME DIFFERENTIALS ACROSS RUSSIAN CITIES[**]

This paper discusses a mechanism underlying the input allocations and income distributions across Russian cities of different ages. The empirical strategy is based on using the extended version of the Glaeser-Gottlieb model to guide the interpretation of regression estimates. The results are in line with previous evidence. Newer cities tend to pay higher real wages, but this is offset by the poor consumption amenities. Their opportunities to pay more are related to their productivity advantages resulting from their higher shares of skilled workforce, and more available natural resources. At the same time, these advantages and disadvantages tend to disappear with time, which gives rise to the income convergence.

JEL Classification: R12, Q32, O13

Keywords: exhaustible resources, urban development, consumer amenities

[*] National Research University Higher School of Economics (Saint-Petersburg, Russia)
e-mail: skorobogatov@mail.ru

[**] The author is grateful to for valuable comments and suggestions to Edward Glaeser, Sergey Izmalkov, Sergey Kokovin, Alexey Korolev, Grigory Kosenok, Jens Suedekum, Jacques François Thisse, Dao-Zhi Zeng, Ives Zenou. All the views and errors are of the author alone.

1. Introduction

Income differences across Russian territory have been paid much attention in the literature.[1] The existing evidence revealed income convergence among Russian regions after 2000, spatial clustering in terms of regional incomes, and the important role of natural resource stocks for the income gaps across the regions (Dolinskaya, 2002; Guriev and Vakulenko, 2012; Herzfeld, 2008; Kholodilin, 2009).

Income disparity dynamics were normally examined using regions as observational units. The researchers mostly dealt with very aggregated and heterogeneous spatial units differing from each other by size, population density and the share of urban population.[2] Cities as observational units are more relevant for a number of research problems. Urban units are much more homogenous with respect to their size and density than regions. At the same time, in the industrialized world cities produce the bulk of national incomes, and focus on income disparity dynamics across cities could allow one to abstact from the agrarian sector and analyze processes specific for the industrialized part of the economy. Despite these advantages, cities are less frequently analysed in terms of their income convergence/divergence. Evidence for Russian cities, in particular, income differentials across them is still scarce. Mikhailova (2011) analyzed city development in Russia as a result of such a historical accident as Stalinist policies to create the labor camps. According to this study, cities that received more initial investment when the camp system was working had a more chance to survive as populated locations. Skorobogatov (2014) documented existence of an inverse link between city age and per capita income in Russia.[3] This relationship is observed not only in Russia, but in other post-Soviet countries.[4] Given the established theory (Krugman, 1991; Davis and Weinstein, 2002), this is a puzzle, and evidence as to the potential process underlying this correlation could be informative for the debate about spatial patterns of input allocation and income distribution. In addition, though age is considered a fundamental characteristic of urban units, there is not yet enough evidence concerning the effect of this charactestic on income and other measures of urban performance. Among few references are Giesen and Suedekum (2012), and Glaeser and Kahn (2001).

Another challenge is the consistent estimation of effects of various forces behind the income differences. For example, population affects income and prices via a number of channels such as agglomeration effects, congestion, and available land for construction, while high income attracts more people and increases housing prices, which means in turn that all the three are endogenous variables.

This paper considers the forces behind spatial income distribution using dataset on Russian cities. The main question of this paper is why city age is inversely linked with income in Russia. The

[1] When it comes to big countries, income differentials across regions may be more important than those across countries. For evidence concerning these differentials, see, e.g., Acemoglu and Dell, 2010).

[2] In particular, territories of some regions reduce to those of cities as in the case of Moscow and St. Petersburg, while the other regions have territories comparable with countries.

[3] This fact can be related to the results of Mikhailova (2011) in that new cities originated under the Soviet regime, including those within the camp system, might attract more investment that resulted in their higher incomes.

[4] For additional details as to the income differences between cities of different ages across and within regions, see Tables A1 and A2 in Appendix. For evidence concerning the other countries, see Table A4 in Appendix.

related issue is the observed income convergence across the cities of different ages.[5] I test traditional hypotheses as to the data-generating process. The first group of hypotheses suggests that younger cities are more productive. I distinguish between human capital, natural resources, and productivity amenities as fources which potentially underlies productivity advantages of the younger cities. The latter may be more productive because either they attract more skilled workers, or they are relatively resource rich locations, or, finally, they have other productivity advantages such as available non-traded capital. Another group of hypotheses is related to the consumption amenities. Younger cities may be worse places for living in which case employers would have to pay bonus to attract people there.

To distinguish between the potential forces behind spatial income distribution, I use spatial equilibrium approach. This is extensively used to deal with similar problems (e.g., Autor and Dorn, 2013; Beaudry, 2014; Glaeser and Gottlieb, 2009; Moretti 2013). Its main assumption reduces to the no-arbitrage condition in terms of individual utilities from one place to another. Higher income in a place may result from higher productivity, which should attract more people. The latter consequence in turn should increase housing prices. Another reason for higher income may be disamenities in a place. Ultimately, higher income in a place is offset either by higher prices, or disamenities, or both, so that the resultant individual utilities should be equal across space. Using these assumptions Glaeser (2008) and Glaeser and Gottlieb (2009) constructed the model, which makes it possible to determine the mechanism underlying the effect of a variable on income, population, and prices. I extend their model to include natural resources. In the case of Russia adding this endogenous variable is motivated by the fact that relative resource abundance heavily affects the spatial income distribution (e.g., Dolinskaya, 2002; Carluer, 2005; Brown et al., 2008).

The structure of this paper is as follows. The next section outlines the spatial equilibrium model to be used for derivation of the mechanism underlying the relationships between exogenous and endogenous variables. The mechanism should guide the interpretation of the regression results. Section 3 describes the data sources and the most important patterns of the data. Section 4 contains the estimation results and the derived interpretation of the estimates based on the spatial equilibrium conditions outlined above. To check the robustness of these conclusions, I estimated a number of regressions of endogenous variables and proxies for the exogenous ones. Section 5 concludes.

2. Spatial equilibrium for inputs and rewards

One of the opportunities to determine the mechanism underlying the effect of a variable on the endogenous variables is to use a theoretical framework. The theory can suggest the relationship between the vectors of exogenous and endogenous variables $\phi : x^k \to y^k$, while the available data can give the vector $\mathbf{b}$ of the slope estimates of regressions of $\mathbf{y}$ on the variable of interest. Then using the theory-

[5]For evidence, see Table A3 in Appendix.

based functions ϕ and the evidence-based estimates $\mathbf{b}$ one can decompose the effect of the interest variable on the endogenous variables into its partial effects on the exogenous variables λ.

A ready framework for deriving the effects of city age on the exogenous variables is the Glaeser-Gottlieb model (Glaeser, 2008; Glaeser and Gottlieb, 2009). The spatial equilibrium model we use is mostly equivalent to the Glaeser-Gottlieb model. Following the classic paper by Roback (1982) and related literature we assume that spatial patterns of input allocation and their rewards are governed by the no-arbitrage condition for consumers, firms, and developers resulting from the free movement of labour and capital.[6] Both individuals and owners of capital are indifferent as to their location and sector because the former have equal utility, and the latter have zero profit whatever location and sector they choose. Actual spatial differentials in real income are counterbalanced by the differences in consumption amenities, while local productivity advantages in production and construction sectors are offset by differences in wages and housing prices. Like in the Glaeser-Gottlieb model, individual preferences, and production and construction technologies are given by Cobb-Douglas functions. Individual firms and developers face a constant scale effect, but at the level of a location there is a diminishing scale effect due to the fixed nontraded capital.

We extend the Glaeser-Gottlieb model to account for the role of natural resources. The latter are included as an input in the respective production and construction functions.[7] Thus, the functions contain four inputs: labour, traded capital, natural resources, and nontraded capital, among which all but the last are subject to an optimal decision at the local level.

2.1. Consumer preferences and demand for housing

An individual should choose the optimal combination of housing H and composite good (other goods) given the locally specific wage W, price of housing P, and given that the composite good is used as a numeraire. The consumer preferences are given by the following Cobb-Douglas utility function:

$$U = \theta(W - PH)^{1-\alpha} H^{\alpha}$$

where θ denotes amenity level in a place of residence.

FOC for housing gives an individual demand for housing $H = \frac{\alpha W}{P}$, whereof one has the indirect utility function:

[6]The potential problem with this assumption is related to the fact that under the Soviet rule the spatial allocation of labour was guided by the state considerations, rather than private interests. However, there is evidence that after the collapse of the planned economy market signals heavily impacted migration (Andrienko and Guriev, 2004). An example of this tendency is the depopulation of remote regions after the state stopped inducing people to stay there (Heleniak, 1999). Though there are still factors that impede migration, including the local preferences of individuals (Moretti, 2011), for the quarter of century of the post-Soviet period a lot of people did move to the locations they preferred. See also Markevich and Mikhailova (2013).

[7]We used natural resources as an input distinct from nontraded capital to keep the decreasing return at the location level.

$$\bar{U} = \xi \theta W P^{-\alpha}. \quad (1)$$

where $\xi = \alpha^{\alpha}(1-\alpha)^{(1-\alpha)}$. The indifference condition suggests that every location provides an individual with the equal utility level. As seen in (1) this implies that high income is offset either by high housing price or poor amenities.

The aggregated housing demand is an individual one times the local population N:

$$D = \frac{\alpha W N}{P}. \quad (2)$$

This will be used when solving for equilibrium housing output.

2.2. Production

The composite good is produced using the Cobb-Douglass technology as follows

$$T N^{\beta} K^{\gamma} R^{\delta} Z^{1-\beta-\gamma-\delta}$$

where T is local productivity in the production of consumer goods, and N, K, R, and Z are labour, traded capital, natural resources, and nontraded capital, respectively. As mentioned, the latter, being fixed at the location level, means the firms have a constant scale effect while their locations face a decreasing scale effect. Since one has to solve for optimal output at the location level, the nontraded capital is constant, meaning that local productivity and the stock of nontraded capital comprise the production amenities of a location $TZ^{1-\beta-\gamma-\delta}$.

Labour price and resource price are denoted as W, and μ, respectively. Wages should offset spatial differences in amenities and housing prices. Capital has equal price of unity everywhere. This follows from the free movement of capital and the assumption that capital does not occupy space. The resource price is inversely related to the remaining stock of the respective resource.[8] We assume that, unlike labour and capital, natural resources do not move, so that their price is an exogenous variable, and firms face different resource costs at various locations. The profit equality across space suggests that the input prices and production amenities offset each other.

A firm maximizes profit $TN^{\beta}K^{\gamma}R^{\delta}Z^{1-\beta-\gamma-\delta} - WN - K - R\mu$.

FOCs for labour, capital, and resources give demand for labor and natural resource

$$W = k_1 \left(T N^{-1+\gamma+\beta} R^{\delta} Z^{1-\beta-\gamma-\delta} \right)^{\frac{1}{1-\gamma}}, \quad (3)$$

$$R = k_2 \left(T N^{\beta} Z^{1-\beta-\gamma-\delta} \mu^{-1+\gamma} \right)^{\frac{1}{1-\gamma-\delta}}. \quad (4)$$

[8]This assumption is in line with the exhaustible resource literature originated by the classic paper of Hotelling (1931). The resource price may take the form of a scarcity-related shadow price.

where $k_1 = \beta\gamma^{\frac{\gamma}{1-\gamma}}$ and $k_2 = \delta^{\frac{1-\gamma}{1-\gamma-\delta}}\gamma^{\frac{\gamma}{1-\gamma-\delta}}$. Since, unlike the other inputs, capital is priced equally across space, the demand for capital is eliminated from the system.

2.3. Construction sector and housing supply

The construction sector uses the Cobb-Douglas technology too, but with different parameters

$$AN^{\epsilon}K^{\eta}R^{\nu}L^{1-\epsilon-\eta-\nu}$$

where A is the productivity in the construction sector, and L is nontraded capital,[9] which, again, suggests a constant scale effect at the firm level and a decreasing scale effect at the location level. The constant $AL^{1-\epsilon-\eta-\nu}$ reflects the construction amenities of a place.

As the construction sector pays the same prices for the inputs as those paid by the production sector, equal profit across space and sectors should make input prices in accordance with construction amenities. A developer maximizes profit $PAN^{\epsilon}K^{\eta}R^{\nu}L^{1-\epsilon-\eta-\nu} - WN - K - R\mu$.

Substituting FOCs for the three inputs in the construction function gives the housing supply function:

$$S = \left(\frac{AL^{1-\epsilon-\eta-\nu}P^{\epsilon+\eta+\nu}}{W^{\epsilon}\mu^{\nu}}\right)^{\frac{1}{1-\epsilon-\eta-\nu}}.$$

Using the demand function (2) one has the equilibrium housing price

$$P = k_3\frac{N^{1-\epsilon-\eta-\nu}W^{1-\eta-\nu}\mu^{\nu}}{AL^{1-\epsilon-\eta-\nu}} \tag{5}$$

where constant $k_3 = \alpha^{1-\epsilon-\eta-\nu}$. As a whole, we have the two linear-dependent equilibria for goods and housing, the latter being explicitly given in (5).

2.4. The effect of city age on the exogenous variables

The equations (1), (3)-(5) contain spatial equilibrium conditions for labor, housing, and natural resources, where N, W, P, and R are the endogenous variables and $TZ^{1-\beta-\gamma-\delta}$, $AL^{1-\epsilon-\eta-\nu}$, θ, and μ are the exogenous variables. Thus, we have four equations, four endogenous variables, and four exogenous variables. Taking the logarithms of the system (1), (3)-(5) and solving it for the endogenous variables we obtain linear equations with the coefficient matrix as follows:

$$\mathbf{c} = \frac{1}{k_4}\begin{pmatrix} 1+\alpha(\eta+\nu-1) & \alpha(1-\gamma-\delta) & 1-\gamma-\delta & \delta(\alpha(1-\eta)-1)+\nu(\alpha(\gamma-1)) \\ \alpha(1-\epsilon-\eta-\nu) & \alpha(\beta+\gamma+\delta-1) & \beta+\gamma+\delta-1 & \alpha(\delta(\epsilon+\eta-1)+\nu(1-\beta-\gamma)) \\ 1-\epsilon-\eta-\nu & \beta+\gamma+\delta-1 & \beta(1-\eta-\nu)+\epsilon(\gamma+\delta-1) & \delta(\epsilon+\eta-1)+\nu(1-\beta-\gamma) \\ 1-\alpha\epsilon & \alpha\beta & \beta & \alpha(\epsilon(1-\gamma)+\beta(\eta-1))-(1-\beta-\gamma) \end{pmatrix} \tag{6}$$

[9]In the case of construction this input mostly consists of land. See Glaeser and Gottlieb (2009, p. 993).

where $k_4 = (1-\alpha\epsilon)(1-\gamma-\delta)+\alpha\beta(1-\eta-\nu)-\beta$. The signs of the coefficients in (6) are in line with the theory. The production amenities (the first column of **c**) are to positively affect all the endogenous variables. The construction amenities (the second column of **c**) make construction costs lower, which makes in turn housing price lower (the negative sign of c_{32}). This attracts more population (the positive sign of c_{12}), which decreases wage (c_{22}) and increases the demand for natural resources (c_{42}). The consumption amenities (the third column of **c**) attracts more population (c_{13}), which is offset by lower wage (c_{23}). The effect on housing price is mixed (c_{33}). More people in both sectors make additional demand for housing, while more people in the construction sector create additional housing supply. The latter increases the resource demand (c_{43}). The resource exhaustion expressed in higher resource price (the fourth column of **c**) makes it rational to use the resources less (c_{44}), which decreases the labour demand as well (c_{14}). The effect on wage (c_{24}) and housing price (c_{34}) is mixed, because, on the one hand, housing demand becomes lower, on the other hand, construction costs become higher.

Based on this model one can decompose the effect of city age on any of the endogenous variables into the partial effects related to the exogenous variables as was done in Glaeser (2008) and Glaeser and Gottlieb (2009). To this end, one needs to estimate regressions of the endogenous variables on city age. Solving the system (1), (3)-(5) in logarithms for the exogenous variables and substituting the coefficients into the solution one obtains the relationships between city age and the exogenous variables:[10]

$$\lambda_T = b_N(1-\beta-\gamma)+b_W(1-\gamma)-b_R\delta, \tag{7}$$

$$\lambda_A = b_N(1-\epsilon-\eta)+b_W(1-\eta)-b_P-b_R\nu, \tag{8}$$

$$\lambda_\theta = b_P\alpha-b_W, \tag{9}$$

$$\lambda_\mu = b_N+b_W-b_R. \tag{10}$$

where λ_T, λ_A, λ_θ, and λ_μ are the city age coefficients from the linear functions of the production, construction, and consumption amenities, and the resource price. b_N, b_W, b_P, and b_R are the city age coefficients from the linear regressions of labour force, wage, housing price, and natural resource use.[11] Thus, having specific values of the parameters and substituting the coefficients from the respective regressions of the endogenous variables into (7)-(10) one can calculate the effect of city age on the exogenous variables and, thereby, determine the sources of the negative age-income correlations coupled with the relationships between city age and other endogenous variables.

[10]For more details concerning the derivation of the parameters for the three equation system, see Glaeser (2008, p. 54-55).

[11]The resultant linear combination in (9) is the same as that in Glaeser (2008, p. 55), while those in (7)-(8) differ in notations and using the resource coefficient. The additional coefficient defined in (10) follows from the inclusion of the new input in the production and construction functions.

2.5. The dynamic system

The first differences of equations (1), (3)-(5) are as follows

$$
\begin{aligned}
\frac{P_{t+1}}{P_t} &= \kappa_4 \left[(1+g_\theta)\frac{W_{t+1}}{W_t}\right]^{\frac{1}{\alpha}}, \\
\frac{W_{t+1}}{W_t} &= \kappa_2 \left[\frac{(1+g_T)\left(\frac{R_{t+1}}{R_t}\right)^{\delta}}{\left(\frac{N_{t+1}}{N_t}\right)^{1-\beta-\gamma}}\right]^{\frac{1}{1-\gamma}}, \\
\frac{R_{t+1}}{R_t} &= \kappa_3 \left[\frac{(1+g_T)\left(\frac{N_{t+1}}{N_t}\right)^{\beta}}{(1+g_\mu)^{1-\gamma}}\right]^{\frac{1}{1-\gamma-\delta}}, \\
\frac{N_{t+1}}{N_t} &= \kappa_1 \left[\frac{(1+g_A)\frac{P_{t+1}}{P_t}}{(1+g_\mu)^{\nu}\left(\frac{W_{t+1}}{W_t}\right)^{1-\eta-\nu}}\right]^{\frac{1}{1-\epsilon-\eta-\nu}}
\end{aligned}
\tag{11}
$$

where g_T, g_A, g_θ, g_μ are exogenous growth rates of the location-specific productivity in the production and construction sectors, consumption amenities, and the resource price, respectively. From the standpoint of an individual, a change of income can be offset by a change of current housing cost, rather than housing price. However, assuming the constant expected growth of housing price $\frac{P_{t+1}}{P_t}$ this can serve as a proxy for the growth rate of rental costs.[12]

The respective solution of the first-differenced equation system for the exogenous variables is as follows

$$\lambda_{1+g_T} = b_{1+g_N}(1-\beta-\gamma) + b_{1+g_W}(1-\gamma) - b_{1+g_R}\delta, \tag{12}$$

$$\lambda_{1+g_A} = b_{1+g_N}(1-\epsilon-\eta) + b_{1+g_W}(1-\eta) - b_{1+g_P} - b_{1+g_R}\nu, \tag{13}$$

$$\lambda_{1+g_\theta} = b_{1+g_P}\alpha - b_{1+g_W}, \tag{14}$$

$$\lambda_{1+g_\mu} = b_{1+g_N} + b_{1+g_W} - b_{1+g_R}. \tag{15}$$

where λ_{1+g_T}, λ_{1+g_A}, λ_{1+g_θ}, λ_{1+g_μ}, b_{1+g_N}, b_{1+g_W}, b_{1+g_P}, and b_{1+g_R}, are the coefficients for the growth rates of the respective exogenous and endogenous variables. It is readily seen that the equations (12)-(15) are the same as those for the static model except that the coefficients from the static regressions are replaced by those from the first-differenced ones.

[12] For a detailed discussion of this assumption, see Glaeser (2008, pp. 70-74).

2.6. Spatial equilibrium for input structure

An additional issue is the spatial allocation of various kinds of inputs, such as human capital and natural resources. To address this issue, we extend the model of Glaeser (2008), which distinguishes between skilled and unskilled workers, to add the distinction between exhaustible and reneweable resources with constant elasticity of substitution between them. The latter distinction is based on the assumption that exhaustible resources are more scarce and expensive, which make them more important for local productivity compared with reneweable ones. In addition, we keep the assumptions of the Glaeser model (2008) concerning skilled and unskilled workers. They have different productivities and wages, different reservation utility levels, and different places of living within a location[13] which provide them with different consumption amenities. Finally, the construction sector produces two kinds of housing with different productivity parameters.

The consumer problem is the same as in (1), but for the two kinds of human capital there are two indifference conditions

$$\bar{U}_S = \xi\theta_H W_S P_H^{-\alpha},$$
$$\bar{U}_U = \xi\theta_L W_U P_L^{-\alpha}$$

where indices S and U refer to the skilled and unskilled workers, and indices H and L refer to the high and low qualities of housing and consumption amenities.

The construction sector uses the technology

$$A(\psi N_S^{\sigma} + N_U^{\sigma})^{\frac{\epsilon}{\sigma}} K^{\eta} (R_e^{\varsigma} + R_r^{\varsigma})^{\frac{\nu}{\varsigma}} L^{1-\epsilon-\eta-\nu} \tag{16}$$

where ψ is productivity parameter, σ and ς denote the substitution parameter between skilled labour N_S and unskilled labour N_U, and between exhaustible resources R_e and renewable resources R_r, respectively.

To derive the equilibrium housing price functions, we use the construction function (16), FOCs for labour, capital, and resources, and housing demand functions based on the indifference conditions, as in the derivation of (5). Then, we substitute the respective equilibrium price function into the indifference conditions and take the respective ratio between the two kinds of workers to obtain

$$\omega = \left[\frac{\iota^{\alpha(1-\epsilon-\eta-\nu)}}{\vartheta\Lambda^{\alpha}}\right]^{\frac{1}{1+\alpha(\eta+\nu-1)}} \tag{17}$$

where $\omega = \frac{W_S}{W_U}$, $\iota = \frac{N_S}{N_U}$, $\vartheta = \frac{U_U\theta_H}{U_S\theta_L}$, and $\Lambda = \frac{A_H L_H^{1-\epsilon-\eta-\nu}}{A_L L_L^{1-\epsilon-\eta-\nu}}$.

The respective technology in the production sector is $T(\psi N_S^{\sigma} + N_U^{\sigma})^{\frac{\beta}{\sigma}} K^{\gamma} (R_e^{\varsigma} + R_r^{\varsigma})^{\frac{\delta}{\varsigma}} L^{1-\beta-\gamma-\delta}$. FOCs for the inputs give the following ratios for the two kinds of workers and resources:

[13]Fu and Gabriel (2012) give evidence for this assumption using Chinese data.

$$\omega = \psi \iota^{\sigma-1}, \tag{18}$$

$$\rho = \phi^{\frac{1}{\varsigma-1}}. \tag{19}$$

where $\rho = \frac{R_e}{R_r}$, and $\phi = \frac{\mu_e}{\mu_r}$.

Finally, the ratio of the housing demand functions is as follows:

$$\pi = \frac{\iota^{1-\epsilon-\eta-\nu}\omega^{1-\eta-\nu}}{\Lambda}. \tag{20}$$

where $\pi = \frac{P_H}{P_L}$.

The solution of the system (17)-(20) for the exogenous variables gives the following relationships between the coefficients of the regressions of the exogenous variables on an interest variable and the regression coefficients of the endogenous variables on the interest variable

$$\lambda_\Lambda = b_\iota(1-\epsilon-\eta-\nu) + b_\omega(1-\eta-\nu) - b_\pi, \tag{21}$$

$$\lambda_\psi = b_\omega + b_\iota(1-\sigma), \tag{22}$$

$$\lambda_\vartheta = b_\pi\alpha - b_\omega, \tag{23}$$

$$\lambda_\phi = b_\rho(\varsigma - 1). \tag{24}$$

where λ_Λ, $\lambda_{\psi'}$, λ_ϑ, λ_ϕ, b_ι, b_ω, b_π, and b_ρ are the coefficients for the respective ratios of the exogenous and endogenous variables. This system allows one to readily move to the dynamic version by substituting the coefficients from the respective first-differenced equations, as in the previous subsection.

2.7. Hypotheses

The theory presented lets one test a number of hypotheses concerning the mechanism behind the inverse age-income relationship. In statics, one can distinguish between productivity-related forces and amenity-related ones.

The former group includes hypotheses according to which newer cities:

(a) Feature higher production amenities;

(b) Attract more skilled workforce. This can be the result of bonuses for human capital, the specific construction amenities, or specific consumption amenities from the standpoint of a skilled worker. In other words, skilled workers may be attracted by either relatively higher wages, or relatively more available housing, or relatively more pleasant places for living, or a combination of these relative advantages;

(c) Are more resource-rich;

(d) Feature higher shares of exhaustible resources in their resource stocks;

A location's productivity advantages, whatever their source, should make its income higher. This in turn should attract additional population, which increases housing prices. Thus, productivity-related higher wages in newer cities would be offset by higher housing prices.

The amenity-related hypotheses include those according to which newer cities:

(a) Feature consumption disamenities. To attract people to relatively unpleasant newer cities, firms and developers staying there would have to pay higher wages;

(b) Feature construction disamenities. Poor conditions for construction in newer cities would make their housing price higher, which, again, should be offset by higher wages.

Finally, to explore the relationship between city age and income change, we tested the first difference versions of the same hypotheses.

3. Data

The main body of the data on Russian cities is from the dataset Ekonomika gorodov Rossii (2015). This contains key dependent variables, namely average wage, labor force, and extractive output of energy and non-energy resources. The latter were of key importance for the hypotheses tested as they were used as proxies for the resource use. Data on control variables also comes from this dataset.

The dataset contains the two age variables, but these are not consistent in terms of a city start. To correct these data, I used a number of sources for city age. One of them is Administrativno-territorialnoye delenie (1987). Other source is Lappo (1998).

A number of important variables are available only at the level of regions, in which cases the respective urban characteristics were proxied by the regional data. The data for average temperatures in January and July, consumer and housing prices were taken from Regiony Russii (2015). The prices for higher quality and typical quality housing were taken from the Rosstat (2015). The shares of population with higher education come from Regiony Russii (2015), and then were corrected to exclude rural population with the use of results of Vserossiyskaya Perepis Naseleniya (2002; 2010). For more details concerning the correction, see Appendix C. Wages by the level of education come from Trud and zanyatost' v Rossii (2007; 2015).

The dynamic version of this model suggests the same dependent variables in their first differences. For this end, I used the change of the dependent variables from 2005 to 2013.

Data on geographical coordinates (in decimal degrees) were taken from Bariev (2007) for 992 observations; the remaining 97 items were taken from Internet maps. Summary statistics for three age groups are presented in Table 1.

As seen in the table, younger cities tend to pay higher nominal wages and, despite their consumer and housing prices are also higher, their real wages is still higher too; they pay relatively higher wage to high-qualified workers, and they have higher shares of educated people, though their human capital measured by the relative numbers of college students is lower; they locate to the north and the east, and suffer from colder climate; other proxies of their consumer amenities such as the number of sports facilities or the quality of their natural environment also tend to be relatively poor; their population size and density tend to be lower; they have better economic indicators such as unemployement rate, relative labor force, but their industrial concentration tends to be stronger; finally, they produce less manufactured goods, but their extractive output is much higher. If one relates these summary statistics to the hypotheses to be tested, higher incomes in younger cities are likely to be resulted from poor consumption amenities and/or productivity advantages related to higher level of human capital and/or resource abundance.

4. Results and discussion

4.1. Econometric specifications

As follows from the models presented above, to test the hypotheses, one needs a number of estimates. To obtain the coefficients (7)-(10) and (21)-(24), based on the static model and on distinguishing between the types of human capital and natural resources, respectively, we ran a number of regressions. As dependent variables in the former case we use the logarithms of labour, wage, housing price, and stocks, and in the latter case we use the logarithms of the share of people with higher education, the high-qualified to low-qualified worker wage ratio, the higher quality to typical quality housing price ratio, and the exhaustible to renewable resource extraction ratio. The latter variable, available in the disaggregated urban dataset, is used as a proxy for the relative resource use. All the regressions are estimated controlling for important economic characteristics, which are not immediately related to either of the dependent variables, namely, the logarithms of the shares of firms in all the organizations, the ratio of the workforce to the number of firms, and the unemployment rate. Standard errors are estimated using the robust estimator allowing for clustering at the region level.

To determine the relationships between city age and the change of all these variables, we run the dynamic specifications like those in Glaeser and Gottlieb (2009) except that we use all the available cross-sections between the earliest and latest years. For obtaining the coefficients (12)-(15) we run the panel regressions for 1996–2013, while for the relative measures the time span is 2005–2009. Thus, our specifications are as follows:

$$y_{it} = b_1 year_t + b_2 year_t \times \log(age_i) + b_3 y_{it-1} + \mathbf{byr}_t + \alpha_i + \varepsilon_{it} \tag{25}$$

where y_{it} is the logarithm of a dependent variable among those described above in the ith city for the tth year, $year$ is the trend variable, age_i is the ith city's age, $\mathbf{yr}$ is the vector of year dummies, α_i is the ith city's fixed effect, and ε_{it} is the ith city's idiosincratic characteristics in the tth year. Assuming that the dependent variable lag y_{it-1} is a close covariate of a number of urban characteristics it is

Table 1. Sample statistics for three age groups

	founded before 1957			founded between 1957 and 1972			founded after 1972		
	obs	mean	sd	obs	mean	sd	obs	mean	sd
Age variables									
Based on year of foundation	996	309.0412	279.4189	56	50.1786	4.0096	23	34.9565	6.0863
Based on year of giving city status	996	194.5050	224.1659	56	35.2321	12.1565	23	27.7391	7.0596
The first variable from the dataset	972	214.2973	244.2579	54	39.8519	29.4262	27	58.9259	137.8031
The second variable from the dataset	972	209.4877	242.6956	54	39.5741	29.3512	27	58.9259	137.8031
Average income									
Wage	973	25,089	9,728	51	34,498	16,831	23	43,408	18,111
Real wage	973	2.3450	0.6772	51	3.0184	1.0606	22	3.5534	0.9615
Geographic characteristics									
Latitude divided by 90	996	0.6054	0.0541	56	0.6124	0.0666	37	0.6511	0.0742
Longitude divided by 180	996	0.3016	0.1554	56	0.3479	0.1520	37	0.3490	0.1998
Distance to the nearest station	996	42.8454	223.1495	56	90.0946	284.6995	38	38.8816	91.0136
Distance to the nearest dock	996	44.4024	98.6118	56	43.3946	126.9551	38	19.6316	40.3769
Demographic characteristics									
Population size	977	98,037	445,600	53	45,762	49,336	24	35,568	28,107
Population density	927	11.8649	10.6438	47	10.4234	11.5720	21	5.7190	4.5835
Net migration to population ratio	973	-0.0026	0.0132	53	-0.0073	0.0112	24	-0.0003	0.0355
Labour force to population ratio	928	0.5844	0.0318	49	0.6096	0.0436	24	0.6518	0.0496
College students to population ratio	996	0.0128	0.0276	56	0.0039	0.0089	42	0.0013	0.0041
Social amenities									
Theatre attendances to population ratio	996	0.0001	0.0002	56	0	0.0001	42	0	0.0001
Discharge from stationary sources to area ratio	996	0.0008	0.0023	56	0.0016	0.0030	42	0.0014	0.0074
Economic characteristics									
Firms to all organizations ratio	996	0.2756	0.1186	56	0.2312	0.0890	42	0.1464	0.1433
Unemployment rate	996	0.0096	0.0156	56	0.0073	0.0150	42	0.0046	0.0083
Labour force to firms ratio	928	59.4549	31.5486	49	72.0700	40.4112	24	67.0946	36.8657
Industrial outputs									
Electricity, gas, and water to labour ratio	996	40.3453	130.2874	56	160.2853	408.7728	42	111.3335	328.2192
Manufacturing output to labour ratio	996	209.2253	406.3291	56	198.6823	369.0112	42	48.9284	114.5952
Extractive output to labour ratio	996	53.4801	472.6981	56	437.5279	1,172.8821	42	179.8800	472.8771
Energy to nonenergy resource extraction ratio	996	462,608	4496305	56	4484165	1.6294e+07	49	6435967	1.8900e+07
Regional controls									
Temperature in January	996	-11.4202	7.2782	56	-15.3286	9.9461	38	-14.9316	9.8659
Average price for the consumer basket	996	10,563	1,357	56	11,188	1,749	38	12,207	1,870
Average price for square meter of housing	993	49,359	11,539	56	50,498	10,162	40	57,594	17,381
Higher quality to typical quality housing price ratio	736	1.1125	0.1441	47	1.1053	0.1278	29	1.1733	0.2188
Share of people with higher education	994	0.3379	0.0506	56	0.3431	0.0490	30	0.3464	0.0535
High-qualified to low-qualified worker wage ratio	996	2.2764	0.2204	56	2.3385	0.1553	40	2.4120	0.2970

Note: All variables, but college students (2012) and ratio of energy resource extraction to nonenergy resource extraction (2009), are for 2013. Sources for the age variables are given in subsection A2, the source for other urban characteristics is Ekonomika gorodov Rossii (2016), the sources for regional variables are given in subsection A3.

included as a control variable. These regressions are estimated by the Arellano-Bond estimator with the interaction term instrumented by its lags. Like the respective specifications in Glaeser and Gottlieb (2009), the interaction term is the interest variable.

4.2. Estimation results

Initial estimates come from sample statistics for age groups of the cities (see Table 1). These suggest that, apart from the correlations established above, newer cities are less populated, they tend to pay relatively higher wages to high-qualified workers, and have higher shares of educated people, though their human capital measured by the relative numbers of college students and their population size and density tend to be lower; finally, their resource extraction and relative exhaustible resource extraction are higher. If one relates these sample statistics to the hypotheses to be tested, higher incomes in newer cities are likely to result from productivity advantages related to a higher level of human capital and/or resource abundance.

Panel A of Table 4 contains the regression estimates for the static model. The variable of interest is highly significant in columns 2–4. As follows from the estimates, newer cities feature higher incomes, housing prices and natural resource use. An additional 1% of city age makes the average values of income, housing price, and resource use lower by 0.17%, 0.04%, and 0.27%, respectively. At the same time, age does not make any significant difference in terms of city size.[14]

Panel regression estimates to establish the comparative dynamics of these variables across different ages are presented in Panel A of Table 5. The results are significant for the variable of interest, except for column 3. Now wage and resource variables go in the opposite direction compared with the static estimates. Like the results in Table 3, between 1996 and 2013 income grew more quickly in older cities. The growth of resource use was also higher in older cities, while population growth in older cities was lower. Thus, city age predicts the opposite dynamics for income and population. These dynamics coupled with the preceding static results may result from in-migration to relatively rich new cities, which do not feature the most rapid income growth.[15] As follows from the spatial equilibrium model, these dynamics in older cities should go hand in hand with a change for the worse in their consumption amenities.

Panel B of Table 4 contains the results for the relative measures serving as dependent variables. The significantly negative estimates for the variable of interest suggest that newer cities pay educated people higher relative wages and use relatively more exhaustible resources compared with their use of renewable resources. The spatial equilibrium approach suggests that the higher wages for skilled workers in newer cities may be offset by specific consumption disamenities.

[14]This result is at odds with the sample statistics for Russian cities and the existing evidence concerning the relationship between city age and city size in developed countries (Giesen and Suedekum, 2014; Dobkins and Ioannides, 2001; Michaels et al., 2012). However, the data for Russian cities are in line with these results when using a different measure of city size or a different specification. In particular, the regressions of population size, rather than log population size, and population density on log age give highly significant positive estimates of the coefficients.

[15]Similar dynamics can be observed in other countries, cf. evidence for Norway (Rattsø and Stokke, 2014).

Table 2. City age and the endogenous variables

Panel A	Dependent variable, 2013			
	log labour	log wage	log housing price	log stocks
	(1)	(2)	(3)	(4)
Log city age	-0.0674	-0.1699***	-0.0413***	-0.2742***
	[0.0590]	[0.0257]	[0.0156]	[0.0954]
Observations	930	930	927	921
R^2adj	0.373	0.334	0.121	0.430
Panel B	Dependent variable, 2009			
	$\log(N_S/N_U)$	$\log(W_S/W_U)$	$\log(P_H/P_L)$	$\log(R_e/R_r)$
Log city age	-0.0167	-0.0401***	-0.0004	-0.9367*
	[0.0134]	[0.0098]	[0.0060]	[0.5386]
Observations	927	929	653	929
R^2adj	0.0261	0.147	0.0079	0.0143

Note: Robust standard errors clustered at the region level are in brackets.
The controls include log unemployment rate, log commercial firms to all organizations ratio, and
log laborforce to commercial firms ratio. *** p<0.01, ** p<0.05, * p<0.1

Table 3. City age and change of the endogenous variables

Panel A	Dependent variable, 1996-2013			
	log labour	log wage	log housing price	log stocks
	(1)	(2)	(3)	(4)
$year \times \log(age)$	-0.0007***	0.0019***	-0.0000	0.0041***
	[0.0002]	[0.0002]	[0.0003]	[0.0012]
Observations	17,473	18,769	16,481	15,576
Panel B	Dependent variable, 2005-2009			
	$\log(N_S/N_U)$	$\log(W_S/W_U)$	$\log(P_H/P_L)$	$\log(R_e/R_r)$
$year \times \log(age)$	-0.0021***	0.0018***	0.0003	0.1281***
	[0.0003]	[0.0005]	[0.0009]	[0.0439]
Observations	7,506	3,215	4,756	3,225

Note: The estimates were obtained using the Arellano-Bond estimator.
GMM standard errors robust to heteroscedasticity and clustering on city level are in brackets.
The specifications include the first lag of dependent variable, trend variable, and year fixed effects.
City fixed effects were controlled for by the first differences. *** p<0.01, ** p<0.05, * p<0.1

The comparative dynamics in these relative urban characteristics measured by the respective panel regression estimates are in Panel B of Table 5. Like the results in Panel A of the same table, all the estimates except for those in column 3 are significant and with the same signs, which is also in line with the logic of the model. Based on the estimates one can state that between 2005 and 2009 the relative wage of skilled workers grew more rapidly in older cities, whereas the share of skilled workers in the labour force in older cities grew more slowly. Again, these comparative dynamics may have reflected a change for the worse in older cities in their consumption amenities from the standpoint of skilled workers.

4.3. Theory-based interpretation

These results coupled with the spatial equilibrium model and the parameter estimates let one calculate the effects of city age on the exogenous variables, which may underlie the relationships of interest.

Most parameter values are obtained estimating the production function on the same dataset for Russian cities (for estimation results, see Table A5 in Appendix D). The remaining parameters are borrowed from the literature. The parameter values used in the calculations are in Table 6.[16]

All the calculations are presented in Table 7. The values of column 1 calculated based on the interest results of Table 4 in Panel A and formulas (7)-(10) are the regression coefficients of $\log(T)$, $\log(A)$, $\log(\theta)$, $\log(\mu)$, respectively, on log city age. As follows from the values, newer cities feature higher production and construction amenities, and lower consumption amenities. This means that newer cities impose lower production and construction costs on the firms and developers, but provide their inhabitants with less pleasant living conditions.[17]

In addition to the advantages of newer cities in their production and construction amenities, they enjoy resource abundance. The coefficient for μ means that newer cities feature much lower prices for natural resources, meaning that the latter are much more available for them. What do these imply for the age-wage relationship? The higher productivity in newer cities due to their production amenities and abundant resources should increase their wages. The same effect is made by their consumption disamenities. The productivity-related wage bonus should attract additional population. This in turn will increase housing prices, the effect of which is weakened by the consumption disamenities because the latter should discourage people from living in newer cities. Finally, the construction amenities should decrease housing prices. To sum up, newer cities feature consumption disamenities, but this disadvantage is offset by higher real wages. The nominal wages are higher to attract people to the unpleasant locations, but the higher productivity in the production sector does not fully transmit to higher housing prices because of both the relatively small population, and therefore lower housing demand, and higher productivity in the construction sector, which lowers housing costs.

Column 2 contains the coefficient values for the dynamic model used to explain the income convergence observed for the cities with different ages. The values are mostly of the opposite signs compared to column 1, which implies a convergence in terms of the respective urban characteristics. According to these calculations, during the period observed newer cities faced more rapid growth of their consumption amenities, but slower growth of their construction amenities. This suggests that within any fixed group new cities are no longer such unpleasant places compared with older cities, while the construction efficiency differential changed in favour of older cities. Smaller differences between the cities in their consumption amenities should have resulted in smaller income differentials, though this effect is weakened by the slower growth of construction amenities in newer cities, which should increase their relative housing prices. At the same time, new cities experienced faster growth of their production amenities and resource prices. Thus, the increased differences between the old and new

[16]The only parameter for which we have not found any ready estimate, because of the lack of proper empirical evidence (Growiec and Schumacher, 2006), is ς. For this parameter we assume substitutability among the two kinds of resources, so that $\varsigma \in [0, 1]$, and let this parameter take on the central value of 0.5 within the acceptable range. However, one can readily check that a change of the parameter within the range does not change the main conclusions about the direction of the relationships between city age and the exogenous values.

[17]The latter feature of the new cities is similar to that of their American counterparts. According to Glaeser (2008, p. 65) during the last four decades "many older cities have become more attractive as places to live".

cities in their production amenities should have resulted in higher real income differentials, while the reduced differences in construction and consumption amenities and resource abundance contribute to real income convergence across cities by age.

The results based on the extended model are presented in column 3. The signs and relative values of the coefficients are the same as those in column 1. The coefficients of ψ and ϕ suggest a higher return on human capital and a lower relative price of exhaustible resources in newer cities. Thus, the additional sources of the new cities' productivity advantages include their more efficient use of human capital and relatively more available exhaustible resources. The remaining coefficients reveal the other comparative characteristics of new cities from the standpoint of skilled workers. Construction productivity in the high-quality housing sector is slightly higher, which increases the relative real wage of skilled workers, while their relative consumption amenities are lower. Skilled workers in newer cities should suffer from consumption disamenities, but enjoy higher real wages. The lack of a significant difference in the shares of educated people by city age (Table 7) suggests that the attractive and non-attractive characteristics of the new cities make skilled workers indifferent between newer and older cities.

Column 4 contains the coefficients for the dynamic version of the extended model, which indicate the change of variables underlying the locational decisions of skilled workers. All the values are of the opposite signs compared to column 3. This, again, suggests convergences in the relative urban characteristics. The return on human capital and the relative abundance of exhaustible resources grew more slowly in newer cities, meaning that the respective productivity-related gaps between the cities of different ages shrank with time. The relative consumption amenities and relative price of high-quality housing grew more rapidly in newer cities. Thus, from the skilled workers' perspective, newer cities became better in the living conditions, but more expensive. The negative sign of the skilled workforce dynamics (Panel B of Table 5) suggests that the final effect of these changes in the return on human capital and relative construction and consumption amenities on human capital allocation was in favour of newer cities.

The estimation results and their theory-based interpretation are in favour of some of the hypotheses. In particular, higher wages in newer cities in Russia result from their higher productivity. The productivity advantages of newer cities are related to their production amenities, the availability of resources and a higher share of exhaustible resources. In addition, they feature more efficient use of skilled workers. The productivity-related higher wages of new cities should be transmitted to higher housing prices. However, their higher construction productivity should make their housing more available, which reinforces their population growth. At the same time, newer cities are generally less pleasant places for living, which should discourage in-migration. Thus, there are both productivity- and disamenity-related reasons for higher nominal incomes in new cities.[18] Recall that the coefficients measure the point elasticities of the respective exogenous variables with respect to city age. By the

[18]This regularity is similar to the general tendency marked for developing countries. Resource-rich countries may rapidly industrialize, but their cities feature worse living conditions (Gollin et al., 2016).

Table 4. The parameters

parameter	value	source
α	0.103	The EU in the world (2015, p. 35)
β	0.6946	Table A5 in Appendix D; Kuboniwa (2011, p. 8), Rõõm (2001, p. 10)
γ	0.0715	Table A5 in Appendix D
δ	0.1338	Table A5 in Appendix D
ϵ	0.7416	Table A5 in Appendix D; Serebryakov (2000, p. 157)
η	0.0554	Table A5 in Appendix D
ν	0.103	Table A5 in Appendix D
σ	0.5	Behar (2010, p. 18)
ς	$]0,1[$	Growiec and Schumacher (2006)

Note: The value of 0.5 for the substitution parameter σ is based on the substitution elasticity ε between skilled and unskilled workers of 2, which is borrowed from Behar (2010), and the formula $\sigma = \frac{\varepsilon-1}{\varepsilon}$ (see, e.g., Combes et al. 2008, p. 55).

Table 5. City age and the endogenous variables

λ-s	static (Table 4, Panel A)	dynamic (Table 5, Panel A)	ratios (Table 4, Panel B)	dynamic ratios (Table 5, Panel B)
	(1)	(2)	(3)	(4)
T/ψ	-0.1211	-0.002	-0.0401	0.0008
A/Λ	-0.0909	0.0011	-0.0337	0.0013
θ/ϑ	0.1656	-0.0019	0.0401	-0.0018
μ/ϕ	0.1043	-0.0029	0.4684	-0.0641

Note: The coefficients in columns 1 and 2 for $\log(T)$, $\log(A)$, $\log(\theta)$, $\log(\mu)$ were calculated according to (7)-(10), and (12)-(15). The coefficients in columns 3 and 4 for $\log(\psi)$, $\log(\Lambda)$, $\log(\vartheta)$, $\log(\phi)$ were calculated according to (21)-(24)

absolute values of the coefficients, the most important forces behind the inverse age-wage relationship are the share of exhaustible resources in the resource stocks of new cities (the point elasticity of 0.5), their consumption disamenities (0.17), production amenities (-0.12), and the general resource abundance (0.1).

The most rapid convergent dynamics are observed for the share of exhaustible resources and the general resource stocks as seen in the change in the relative exhaustible resource price (-0.064) and in the change in the resource price (-0.003). These can result from resource exhaustion.[19] Most other urban characteristics, including general and relative consumption and construction amenities, and the return on human capital, show convergence too. Newer cities are better in their construction amenities (-0.09) and in their return to human capital (-0.04), and worse in their consumption amenities, but their advantages and disadvantages became smaller with time. The only characteristic showing divergent dynamics is production amenities. Newer cities tend to have relative advantage in production amenities, and this advantage is growing with time (-0.002).

4.4. Results by subsamples

The same results as those presented in Tables 4, 5, and 7 are obtained for a subsample of resource extracting cities and for the remaining cities. The ultimate calculations for the subsamples, like those in Table 7, are in Table A6 in Appendix E. All the static results in columns 1 and 3 have the same sign for both subsamples. Among the resource extracting cities the effects of city age on the resource

[19]This is in line with Hoteling's model (1931).

stocks and the share of exhaustible resources in resource stocks are much stronger.[20] The age effects on consumption amenities and the specific consumption amenities for the skilled workers are also stronger for resource extracting cities.

The static results for production and construction amenities, the return on human capital and the relative construction amenities have the same sign and similar values among the subsamples. The dynamic regressions display divergence in production amenities and convergence in resource abundance for both groups with a higher speed for these tendencies in resource extracting cities.

For the other characteristics resource extracting cities differ from the remaining cities. Resource extracting cities display convergence in general and specific construction and consumption amenities, and in return on skills, while the remaining cities show divergence in both construction amenity variables and in return on skills, and no tendency in consumption amenities. Finally, resource extracting cities, unlike the remaining cities, show strong convergence in their exhaustible resource share.

All the effects for resource extracting cities shown in Panel A of Table 5A are the same by sign as those in Table 7 and stronger by the absolute values for the resource, consumption amenity, and human capital variables. As seen in Panel B, the remaining cities display either the same effects by their sign to a much lesser extent or display the opposite effects as for the divergence in general and specific construction amenities and the return on skills.

To sum up, the separate results for the subsamples show that the most substantial relationships between city age and the resource and consumption amenity variables are shown by resource extracting cities. As for the other effects, the most important difference between the groups is that resource extracting cities display convergence for most production and construction amenity variables, while the remaining cities show divergence for all these variables.

4.5. Robustness check

Because of the lack of proper data on the exogenous variables we heavily rely on theory when testing the hypotheses. To make sure the theory and the conclusions fit the data, we have estimated a number of regressions using proxies for the exogenous variables. If the theory-based conclusions are correct the regressions of the proxies should be in line with them. Like Glaeser (2008), we use temperature in January as a proxy for consumption amenities. The other two proxies used are the quality of the natural environment and theatre availability.[21] These are measured by the relative discharge from stationary sources and theatre attendance per capita, respectively. A warmer climate, lower relative discharge, and higher theatre availability indicate better consumption amenities. As an inverse proxy for the resource price we use log extractive output per capita. In this case we suppose that high values for the extractive output measure are consistent with a low price of the natural resources.

[20] In the latter case we have 1.472 versus zero because of the resource data we used. Zero just reflects the lack of values for the remaining cities.

[21] A similar proxy for amenities, restaurant availability, was used in Glaeser, Kolko, Saiz (2001).

The model suggests (e.g., (6)) that consumption amenities attract population and increase the resource demand, but are normally offset by lower real income. To check how much the model fits the data, we have run the regressions of the endogenous variables on the proxies for consumption amenities and the resource price. The regressions of log population, log real wage, and log stocks are estimated controlling for the log real wage, log population, and the two production variables, respectively. The results are in Table 8. Temperature in January and theatre availability are consistent with bigger cities, while the temperature and the quality of the natural environment suggest lower real wages, meaning that, in line with the theory, more pleasant cities are more populated (column 1) and pay less in real units (column 2). The model also predicts a positive association between natural resource use and consumption amenities. Among the three proxies for the amenities the discharge ratio and the theatre attendance variables correlate with the stocks in line with the model, while temperature is insignificant.

The regressions of consumption amenity measures on log city age let one check the robustness of the conclusions. According to the conclusions, newer cities are relatively resource-rich, but suffer from poor consumption amenities. If our proxies are relevant, this suggests that newer cities should be colder, dirtier, and feature fewer theatre attendances, while their resource extraction measure should be higher. According to the sample statistics (Table 1), newer cities are located to the north and the east, and suffer from a colder climate; other proxies for their consumption amenities such as theatre attendance or the quality of their natural environment also tend to be relatively poor. The respective regressions are fully in line with these initial observations and the conclusions about the positive relationship between city age and consumption amenities. As seen in Table 9, newer cities tend to be colder, dirtier, and to have fewer theatre performances (columns 1–3). At the same time, the negative relationship between city age and resource extraction indicates the relative resource abundance in new cities (column 4).

Finally, according to our conclusions, the dynamics of consumption amenities and the resource abundance measures show a convergence across the cities by age. To check this result using the proxies for consumption amenities and resource abundance, we have run panel regressions (25). The latter include dependent variable lags, and city and year fixed effects. The differences in dynamics across cities of different ages are captured by the interaction term between the trend variable and log age. The results are presented in Table 10. The signs of interaction term are in line with the conclusions in all the specifications. During 2005–2013 the dynamics of winter temperatures, the quality of environment and the theatre attendance were more favourable in newer cities. One of the most important factors behind the income convergence across the cities of different ages was the convergence in terms of the available resources. Again, if our resource proxy is relevant and the conclusion about the dynamics fits the data, one should observe more favourable dynamics in resource extraction in older cities. The respective estimation results are in column 4. In line with the theory-based conclusion, the general dynamics in resource extraction was more favourable in older cities.

To sum up, the robustness check confirmed that both the model used and the model-based conclusions fit the data. According to the model, consumption amenities should have correlated positively with city size and resource use, and negatively with real wages. The conclusions based on the static model predict that newer cities should be less pleasant places to live, but richer in resources. The calculations based on the dynamic model predict that income convergence should have resulted from the respective convergence in terms of resource abundance and consumption amenities. The consumption amenities correlate with population size, real wage, and the resource variable in the way suggested by the model. In line with the conclusions, newer cities are richer in resources, but colder and dirtier. This should make their real wages higher for productivity- and amenity-related reasons. Finally, income convergence across cities of different ages goes hand in hand with the respective dynamics in these comparative advantages in productivity and amenities.

Table 6. Exogenous and endogenous variables, 2013

	Log population	Log real wage	Log stocks
	(1)	(2)	(3)
Temperature in January	0.0319***	-0.0118***	0.0011
	[0.0092]	[0.0020]	[0.0083]
Discharge ratio	10.2720	8.2374***	-20.6745*
	[14.8760]	[1.8914]	[12.3594]
Theatre attendances ratio	3,139.1239***	-42.7212	1,045.1501***
	[556.0405]	[38.1028]	[274.2881]
Log extractive output per capita	0.0394**	0.0274***	0.1027***
	[0.0189]	[0.0069]	[0.0366]
Log real wage	1.5632***		
	[0.2278]		
Log population		0.0848***	
		[0.0113]	
Log manufacturing output			0.5134***
			[0.0329]
Log electricity, gas, and water output			0.3723***
			[0.0415]
Observations	1,046	1,046	853
R^2 adj.	0.399	0.332	0.725

Note: Robust standard errors clustered at the region level are in brackets, *** p<0.01, ** p<0.05, * p<0.1

Table 7. Age and consumption amenities, 2013

	Temperature in January	Discharge ratio	Theatre attendances ratio	Log extractive output per capita
	(1)	(2)	(3)	(4)
Log age	1.7453***	-0.0003***	0.0000***	-0.5123***
	[0.2613]	[0.0001]	[0.0000]	[0.0608]
Observations	1,075	1,075	1,075	1,075
R^2	0.0398	0.0107	0.0190	0.0619

Note: Robust standard errors clustered at the region level are in brackets, *** p<0.01, ** p<0.05, * p<0.1

Table 8. City age and consumer disamenities, 2005-2013

	Temperature in January	Discharge ratio	Theatre attendances ratio	Log extractive output per capita
	(1)	(2)	(3)	(4)
$year \times \log(age)$	-0.015929*	0.000051***	-0.000001***	0.009733***
	[0.009091]	[0.000015]	[0.000000]	[0.003466]
Observations	9,602	9,675	9,675	7,525

Note: The estimates were obtained using the Arellano-Bond estimator.
GMM standard errors robust to heteroscedasticity and clustering on city level are in brackets.
The specifications include the first lag of dependent variable, trend variable, and year fixed effects.
City fixed effects were controlled for by the first differences.
*** p<0.01, ** p<0.05, * p<0.1

5. Conclusion

This paper deals with a strong statistical regularity that seems to be at odds with the established theory in spatial economics. In Russia and other post-Soviet countries newer cities are substantially richer, despite their smaller population sizes and thereby weaker agglomeration forces. To determine the underlying relationships between city age and other urban characteristics, we use an extended version of the Glaeser-Gottlieb spatial equilibrium model. The model assumes the no-arbitrage condition for population and firms when it comes to their locational decisions. Based on these assumptions the equilibrium conditions in consumer and housing markets were derived where population, wage, housing price, and resource use are endogenously determined. Another version of the model establishes the equilibrium conditions for the share of skilled workers in the local labour force and the share of exhaustible resources in local resource stocks. Both versions are used to determine the dynamics of these variables. The model coupled with a regression analysis allows us to determine the characteristics of the new cities which make them better off in terms of their average wages, and make conclusions concerning the impact of city age on other urban characteristics.

According to the conclusions, age is linked with both productivity- and amenity-related urban characteristics. New cities are more productive. Productivity advantages of new cities result from their production amenities, higher return on skills, more abundant natural resources and higher shares of exhaustible resources in their resource stocks. At the same time, new cities are less pleasant places to live. Higher real wages in new cities are both the reward for their higher productivity and compensation for their poor consumption amenities. The differences between the cities of different ages in their production- and amenity-related characteristics tend to diminish with time. There is a convergence in both income and the underlying urban characteristics among cities by age.

These results for resource use reveal a particular pattern of Russian urban development. Specifically, new cities occur at sites that have rich deposits of valuable resources. The new territory provides new cities with temporary advantages with respect to their resource endowments and ensuing incomes, but feature poor consumption amenities. As time goes by, the resources become depleted and the respective advantages vanish. At the same time, as the territory is developed their disadvantage in consumption amenities also tend to diminish. From a broader perspective, this pattern corresponds to the way the Russian population and state used their vast territory throughout history, moving on from resource depleted areas to new richer areas. While this pattern is at work in Russia, it can be helpful in future studies when examining the spatial development of other resource-rich countries and regions. In particular, a similar mechanism may underlie the negative age-wage relationships in other post-Soviet countries. This pattern suggests the potentially important role of exhaustible resources in the changing regional differences in economic activities and incomes. Over longer historical intervals this may imply a potential effect of technological progress on the dynamics of spatial income distribution via changing relative values of various natural resources.

Appendix A. The inverse city age income relationship in Russia

The inverse city age income relationship is highly robust to change of definition of city age, inclusion of controls, and the sample. The results of estimation of a number of specifications are presented in Table A1.

Table A1. City age and log wage, Russia

	first age variable				second age variable			
	nomin., 2013	real, 2013	controls, 2013	2001	nomin., 2013	real, 2013	controls, 2013	2001
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log age based on:								
foundation year	-0.1282***	-0.1009***	-0.0587***	-0.0869***				
	[0.0127]	[0.0098]	[0.0078]	[0.0169]				
year of giving					-0.0839***	-0.0662***	-0.0451***	-0.0841***
city status					[0.0107]	[0.0085]	[0.0068]	[0.0145]
Regional dummies	No	No	Yes	Yes	No	No	Yes	Yes
p-value of F-stat								
for controls:								
Geographic			0.0036***	0.0005***			0.0080***	0.0006***
Demographic			0.0000***	0.0000***			0.0000***	0.0000***
Observations	1,046	1,046	933	963	1,046	1,046	933	963
R^2 adj.	0.105	0.103	0.8216	0.6906	0.0563	0.0555	0.8176	0.6933

Note: Robust standard errors clustered at the region level are in brackets. p-value of F-statistics is for the geographic and demographic characteristics. Geographical controls include latitude and longitude, log distances to railroads and docks, and dummy for the status of regional administrative center. Demographic controls include log population size and log density, net migration per capita, and log labor per capita.
The specifications for 2001 include also the student ratio. *** p<0.01, ** p<0.05, * p<0.1

The estimation results in Table A1 imply that the inverse city age income relationship is robust to inclusion of the full list of regional dummies. To check this relationship at the level of distinct regions, I run a number of respective regressions. The results of presented in Table A2.

Table A2. City age and real wage, distinct regions

regions	Altaiskyi	Volgogradskaya	Vologodskaya	Zabaykalsky	Kaluzhskaya	Krasnoyarsky
	(1)	(2)	(3)	(4)	(5)	(6)
Log age	-0.155*	-0.0690*	-0.243*	-0.259*	-0.0800***	-0.0758**
	(0.0617)	(0.0380)	(0.127)	(0.105)	(0.0144)	(0.0258)
Observations	10	19	15	10	19	21
p-value of F-statistics	0.00310	9.82e-07	0.000437	0.00489	0	9.88e-11
R^2 adj.	0.749	0.137	0.455	0.710	0.880	0.909
regions	Kurskaya	Moskovskaya	Nizhegorodskaya	Orenburgskaya	Komi	Ryazanskaya
	(7)	(8)	(9)	(10)	(11)	(12)
Log age	-0.221**	-0.0736***	-0.101**	-0.279**	-1.029***	-0.101**
	(0.0701)	(0.0273)	(0.0381)	(0.104)	(0.0451)	(0.0342)
Observations	10	77	27	12	10	12
p-value of F-statistics	0.0395	0	0	3.73e-05	4.16e-05	0.000350
R^2 adj.	0.633	0.351	0.573	0.504	0.981	0.490

Note: Robust standard errors obtained by the sandwitch estimator are in brackets. *** p<0.01, ** p<0.05, * p<0.1.
p-value of F-statistics is for the controls. The controls include log population, dummy for administrative centre, and latitude and longitude.

Table A3. Nominal and real wages and city age in dynamics, 1991-2013

dependent variable	log wage			
	nominal	real	nominal	real
	(1)	(2)	(3)	(4)
$year$	0.1564***	0.0576***	0.1601***	0.0618***
	[0.0027]	[0.0047]	[0.0022]	[0.0037]
$year \times \log(age_1)$	0.0031***	0.0028***		
	[0.0005]	[0.0008]		
$year \times \log(age_2)$			0.0027***	0.0023***
			[0.0004]	[0.0007]
Observations	23,818	13,683	23,818	13,683
R^2 within	0.9919	0.8369	0.9919	0.8367

Note: Robust standard errors obtained by the sandwitch estimator are in brackets.
City and year fixed effects are included.
age_1 is based on foundation year, age_2 is based on the year of giving city status.
*** p<0.01, ** p<0.05, * p<0.1

Finally, one can consider city age-income relationship in dynamics to understand how income changes depending on city age. For this end, I estimated panel regressions of log wage on year and its interaction with log city age. The sign of the interaction term indicates the direction of relative income change. Positive sign would indicate income convergence, meaning that younger cities, though more rich, have slower income growth compared with older ones and vica versa. To check the robustness of the results, the regressions were estimated with individual and year fixed effects, for nominal and real wages, and for the two age variables. The results are presented in Table A3. The interaction terms of our interest are highly significant, and their positive signs indicate the convergence.

Appendix B. The relationship in other post-Soviet countries

Is the inverse relationship between city age and income exclusively Russain anomaly? Available data on post-Soviet countries enable one to check if there are similar correlations elsewhere. I estimated similar specifications for a number of countries given the data on city age, average wage, and population size or density. The data were borrowed from the official websites of the statistical agencies. In most cases it was possible to use municipal statistics containing data for the important cities of the respective region and its subregions. Thus, for the important cities I had ready data, while for small towns I used the data for the subregions for which they serve as their administrative centers. The results are presented in Table A4.

Table A4. Other post-Soviet countries

	Ukraine, 2013			Belarus, 2014			Lithuania, 2010	
	simple	+ log pop.	+ regions	simple	+ log pop.	+ regions	simple	+ log pop.
Log age	-0.1096***	-0.1048***	-0.0893***	-0.0360	-0.0523**	-0.0622***	-0.0929	-0.2352**
	[0.0146]	[0.0138]	[0.0200]	[0.0329]	[0.0240]	[0.0206]	[0.0869]	[0.0707]
Log pop.		0.0648***	0.0534***		0.0742***	0.0694***		0.2390**
		[0.0076]	[0.0064]		[0.0098]	[0.0094]		[0.0915]
Obs.	458	458	458	112	112	112	10	10
R^2 adj.	0.173	0.268	0.372	0.016	0.383	0.483	0.089	0.471
	Kazakhstan, 2013			Tadzhikistan, 2013				
	simple	+ log pop.	+ regions	simple	+ log pop.	+ regions		
Log age	-0.0219	-0.1424**	-0.1264**	-0.1242*	-0.1539**	-0.0909		
	[0.0512]	[0.0580]	[0.0502]	[0.0645]	[0.0561]	[0.0988]		
Log pop.		0.1120***	0.0993***		0.1209	0.0479		
		[0.0287]	[0.0242]		[0.0704]	[0.2530]		
Obs.	87	87	87	17	17	17		
R^2 adj.	0.002	0.173	0.610	0.220	0.202	0.063		
	Azerbaidzhan, 2011			Karabakh, 2011			Moldova, 2014	
	simple	+ log pop.	+ regions	simple			simple	+ log pop.
Log age	0.0138	-0.0367***	-0.0349***	-0.0341				
	[0.0240]	[0.0114]	[0.0124]	[0.0182]				
Log pop.		0.1054***	0.0582***					
		[0.0339]	[0.0178]					
Age							-0.0013*	-0.0018**
							[0.0007]	[0.0006]
Age sq.							0.0000**	0.0000***
							[0.0000]	[0.0000]
Log dens.								0.0386
								[0.0226]
Obs.	60	60	60	9			32	26
R^2 adj.	0.016	0.331	0.592	0.340			0.104	0.322

Note: Robust standard errors obtained by the sandwitch estimator are in brackets. *** p<0.01, ** p<0.05, * p<0.1. The data sources are given in Appendix A.6. All the income variables are annual ones, except for Moldova for which the income variable is a monthly one for September.

As seen in the table similar correlations are observed in, at least, eight other countries. Among the other countries Ukraine is the most similar to Russia in terms of the values, statistical significance, and robustness of the inverse city age-income relationship. The latter is highly significant both in the simple specification and with inclusion of log population and all the regional dummies. The second most similar relationship to Russian one is displayed by Belarus. It is worth noting that these two countries are also the most urbanized among the others if it is measured in the numbers of cities. The remaining countries in the table display the significant inverse city age-income relationships depending on a specifications. All of them have the inverse relationship after inclusion of log population, which is explained by the latter being positively correlated both with city age and income. Where it is possible

I include all the regional dummies, and in all these cases, but Tadzhikistan, the inverse relationship is robust to this inclusion. The relationship is nonlinear for all the countries, which is specified in the logarithms, but in the case of Moldova the nonlinearity turns out a more spectacular, so that the significant inverse relationship is observed only in the quadratic, rather that logarithimic, specification.

As a whole, the negative city age-income correlation holds for most post-Soviet countries. Though these countries have common Soviet legacies, they dramatically differ from each other in their territorial location, political regime, and the economy. Thus, one can still argue that the inverse city age-income correlations feature diverse countries, and one can suppose existence of similar correlations elsewhere in the globe.

Appendix C. Correction of the share of urban population with higher education

The data for average temperature in January, and consumer and housing prices are taken from Regiony Russii (2015). The prices for higher quality and typical quality housing are taken from Srednaya tsena (2016). Wages by level of education come from Trud i zanyatost v Rossii (2007; 2009; 2011; 2013; 2015). The shares of population with higher education come from Regiony Russii (2015). These are corrected to exclude rural population with the use of results of the population censuses of 2002 and 2010 (Vserossiyskaya Perepis Naseleniya, 2002; 2010). The dataset Regiony Rossii (2015) contains the share of total population with higher education $higher_{all2005}$. As this paper considers only urban units, we need to correct them to obtain this variable for urban population. To make the correction, we use the following formula:

$$share_{higher-educ2005} = \frac{higher_{all2005} - (1-w) \times higher_{rural2002} \times \frac{higher_{all2005}}{higher_{all2002}}}{w}$$

where $higher_{rural2002}$ is the share of rural population with higher education and PhD according to Census of 2002 (Vserossiyskaya Perepis Naseleniya, 2002), and w is the share of urban population in 2005 according to Regiony Rossii (2015). The numbers in Census of 2002 (Vserossiyskaya Perepis Naseleniya, 2002) and Regiony Rossii (2015) differ because of the difference in data between 2002 and 2005, and the difference in the results of the respective surveys. Assuming that this difference is constant across regions, we correct this difference with the share of the population with higher education according to Census of 2002 (Vserossiyskaya Perepis Naseleniya, 2002) $higher_{all2002}$. The same correction is made for the share of the population with higher education in 2006 in line with Census of 2002 (Vserossiyskaya Perepis Naseleniya, 2002) and for 2007-2013 using Census of 2010 (Vserossiyskaya Perepis Naseleniya, 2010).

Appendix D. Production function

Table A5. Production function estimation

dependent variable	Log output	Log construction
	(1)	(2)
Log labour force	0.6946***	0.7416***
	[0.0097]	[0.0260]
Log fixed capital stocks	0.0715***	0.0554**
	[0.0093]	[0.0229]
Log resource stocks	0.1338***	0.1030***
	[0.0109]	[0.0279]
Observations	9,106	7,191
p-value of F-test of equality to 0.9	0.3756	0.3522
R-squared within	0.4424	0.1749

Note: Robust standard errors obtained by the sandwitch estimator are in brackets.
*** p<0.01, ** p<0.05, * p<0.1. City and year fixed effects are included. The presented estimates are for the constrained regressions, the constraint being that the sum of parameters equals to 0.9.

Appendix E. City age and the endogenous variables by subsamples

Table A6. City age and the endogenous variables

λ-s	(1)	(2)	(3)	(4)
Panel A: resource cities				
T/ψ	-0.1029	-0.0019	-0.049	0.0013
$A/A(r)$	-0.0669	0.0022	-0.0412	0.0017
$\theta/\theta(r)$	0.1943	-0.0025	0.049	-0.0023
$\mu/\mu(r)$	0.1523	-0.003	1.472	-0.1187
Panel B: non-resource cities				
T/ψ	-0.1179	-0.001	-0.0469	-0.0011
$A/A(r)$	-0.092	-0.0011	-0.0284	-0.0002
$\theta/\theta(r)$	0.1345	0	0.0296	0
$\mu/\mu(r)$	0.0349	-0.0013	0	0

Note: See Table 7.

References

[1] Acemoglu, D., Dell, M., 2010. Productivity differences between and within countries. American Economic Journal: Macroeconomics, 2 (1), 169-188.

[2] Andrienko, Yury, Guriev, Sergey M., 2004. Determinants of interregional mobility in Russia. Economics of Transition 12 (1), 1-27.

[3] Autor, David H., Dorn, David, 2013. The growth of low-skill service jobs and the polarization of the US labor market. American Economic Review 103 (5), 1553-1597.

[4] Bariev, I. K. (2007) Koordinaty naselennykh punktov, chasovye poyasa i ismenenia ischisleniya vremeni. Spravochnik. Konek, Moscow.

[5] Beaudry, Paul, Green, David A., Sand, Benjamin M., 2014. Spatial equilibrium with unemployment and wage bargaining. Theory and estimation. Journal of Urban Economics 79 (1), 2-19.

[6] Behar A., 2010. The elasticity of substitution between skilled and unskilled labor in developing countries is about 2. International Monetary Fund. Selected works.

[7] Brown, D., Fay, M., Felkner, J., Lall, S. V. and Wang, H. G., 2008. Death of distance? Economic implications of infrastructure improvement in Russia. European Investment Bank, Economic and Financial Studies, EIB Papers No. 10/2008.

[8] Carluer, F., 2005. Dynamics of Russian regional clubs: the time of divergence. Regional Studies 39 (6), 713-726.

[9] Dobkins, L. Harris, Yannis, M. Ioannides (2001) Spatial interactions among U.S. Cities: 1900-1990. Regional Science and Urban Economics 31 (6), 701-731.

[10] Dolinskaya, Irina, 2002. Transition and regional inequality in Russia: reorganization or procrastination? IMF Working Paper 02/169.

[11] Ekonomika gorodov Rossii, 2016. GMZ Rosstata. <http://multistat.ru/?menu_id=9310004> (accessed February, 2016).

[12] Fu, Yuming, Gabriel, Stuart A., 2012. Labor migration, human capital agglomeration and regional development in China. Regional Science and Urban Economics 42, 473-484.

[13] Giesen, Kristian, Suedekum, Jens, 2014. City age and city size. European Economic Review 71, 193-208.

[14] Glaeser, Edward L., 2008. Cities, agglomeration, and spatial equilibrium. Oxford University Press, Oxford and New York.

[15] Glaeser, Edward L., Kahn, Matthew E., 2001. Decentralized employment and the transformation of the American city. NBER Working Paper No. 8117.

[16] Glaeser, Edward L., Gottlieb, Joshua D., 2009. The wealth of cities: agglomeration economies and spatial equilibrium in the United States. Journal of Economic Literature 47 (4), 983-1028.

[17] Glaeser, Edward, Kolko, Jed, Saiz, Albert, 2001. Consumer city. Journal of Economic Geography 1, 27-50.

[18] Gollin Douglas, Jedwab, Remi, Vollrath, Dietrich, 2016. Urbanization with and without industrialization. Journal of Economic Growth 21 (1), 35-70.

[19] Growiec, J., Schumacher, I., 2006. On technical change in the elasticities of resource inputs // CORE Discussion Paper 2006/63.

[20] Guriev, Sergei, Vakulenko, Elena, 2012. Convergence between Russian regions. CEFIR/NES Working Paper No 180.

[21] Heleniak, Timothy, 1999. Out-migration and depopulation of the Russian North during the 1990s. Post-Soviet Geography and Economics 40 (3), 155-205.

[22] Herzfeld, Thomas, 2008. Inter-regional output distribution: a comparison of Russian and Chinese experience. Post-Communist Economies 20 (4), 431-447.

[23] Hotelling, Harold, 1931. The Economics of exhaustible resources. Journal of Political Economy 39 (2), 137-175.

[24] Kholodilin, Konstantin A., Oshchepkov, Alexei, Siliverstovs, Boris, 2009. The Russian regional convergence process: where does it go? KOF Working Paper No. 216.

[25] Kuboniwa M., 2011. Russian growth path and TFP changes in light of the estimation of production function using quarterly data. RRC Working Paper Series. No. 30.

[26] Lappo, Georgy M. (ed.), 1998. Goroda Rossii. Enziklopediya. Bolshaya Rossiyskaya Enziklopediya, Moscow.

[27] Markevich, Andrei, Mikhailova, Tatiana N., 2013. Economic geography of Russia. The Oxford Handbook of the Russian Economy. Alexeev, Michael, Weber, Shlomo (eds.). Oxford University Press, Oxford.

[28] Michaels Guy, Rauch, Ferdinand, Redding, Stephen J., 2012. Urbanization and structural transformation. Quarterly Journal of Economics 127 (2), 535-586.

[29] Mikhailova, Tatiana, 2011. Looking for multiple equilibria in Russian urban system. Working Paper No10/08E.

[30] Moretti, Enrico, 2011. Local labor markets. Handbook of Labor Economics. Vol. 4b, 1237-1313.

[31] Moretti, Enrico, 2013. Real wage inequality. American Economic Journal: Applied Economics 5 (1), 65-103.

[32] Rattsø, Jørn, Stokke, Hildengunn E., 2014. Population divergence and income convergence: regional distribution dynamics for Norway. Regional Studies 48 (11), 1884-1895.

[33] Regiony Russii: Sozialno-ekonomicheskiye pokazateli, 2015. Rosstat, Moscow.

[34] Roback, Jennifer, 1982. Wages, rents, and the quality of life. Journal of Political Economy 90 (6), 1257-78.

[35] Rõõm, Marit, 2001. Potential output estimates for Central and East European countries using production function method. Eesti Pank.

[36] Serebryakov, Georgiy, 2000. Russian economy: structural change and econometric perspective. In Welfens, Paul J. J., Gavrilenkov, Evgeny. (eds.) Restructuring, stabilizing and modernizing the new Russia. Economic and institutional issues. Springer, New York.

[37] Srednaya tsena 1 kv. m. obshchey ploshchadi kvartir na rynke zhilya. Rosstat, 2016. <http://www.gks.ru/dbscripts/cbsd/dbinet.cgi?pl=1905001> (accessed February, 2016).

[38] SSSR. Administrativno-territorialnoye delenie soyuznykh respublik, 1987. Izvestiya Sovetov Narodnych Deputatov, Moscow.

[39] The EU in the world - 2015 edition. Luxembourg: Eurostat, Publications Office of the European Union, 2015.

[40] Trud i zanyatost v Rossii, 2007. Rosstat, Moscow.

[41] Trud i zanyatost v Rossii, 2009. Rosstat, Moscow.

[42] Trud i zanyatost v Rossii, 2011. Rosstat, Moscow.

[43] Trud i zanyatost v Rossii, 2013. Rosstat, Moscow.

[44] Trud i zanyatost v Rossii, 2015. Rosstat, Moscow.

[45] Vserossiyskaya Perepis Naseleniya, 2002. <http://www.perepis2002.ru/ct/doc/TOM_03_03.xls> (accessed February, 2016).

[46] Vserossiyskaya Perepis Naseleniya, 2010. <http://www.gks.ru/free_doc/new_site/perepis2010/croc/Docu 03-01.xlsx> (accessed February, 2016).

Alexander S. Skorobogatov
National Research University Higher School of Economics
e-mail: skorobogat@mail.ru