Why are newer cities better off in Russia? A spatial equilibrium approach*

Alexander S. Skorobogatov**

This paper documents the negative relationship between the age of cities and their average wages in Russia and a number of post-Soviet countries. To determine a mechanism behind this relationship, a spatial equilibrium model was developed. The empirical strategy is reduced to using the model to guide the interpretation of regression estimates. According to the results, higher real wages in newer cities reflect both their disadvantages as places for living and their production advantages. The latter are related to their production and construction amenities, higher shares of skilled workers, and more available natural resources. At the same time, these advantages and disadvantages tend to disappear over time, which gives rise to income convergence.

JEL Classification: R13, D51, P25
Keywords: exhaustible resources, urban development, consumer amenities

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

** The author is grateful for valuable comments to Sergey Izmalkov, Sergey Kokovin, Alexey Korolev, Grigory Kosenok, Jens Suedekum, Jacques François Thisse, Dao-Zhi Zeng, Ives Zenou. Special thanks go to the Co-Editor of this journal, Edward Glaeser for the important suggestions based on which the author has deeply revised the paper. All the views and errors are of the author alone.

Why are newer cities better off in Russia? A spatial equilibrium approach

Alexander S. Skorobogatov

National Research University Higher School of Economics

St. Petersburg, Russia

Abstract

This paper documents the negative relationship between the age of cities and their average wages in Russia and a number of post-Soviet countries. To determine a mechanism behind this relationship, a spatial equilibrium model was developed. The empirical strategy is reduced to using the model to guide the interpretation of regression estimates. According to the results, higher real wages in newer cities reflect both their disadvantages as places for living and their production advantages. The latter are related to their production and construction amenities, higher shares of skilled workers, and more available natural resources. At the same time, these advantages and disadvantages tend to disappear over time, which gives rise to income convergence.

1. Introduction

Russian territories display a rather unexpected relationship between city age and wages. While the ancient and famous Russian cities are relatively poor, those founded just a few decades ago are relatively rich and are frequently the most flourishing cities. The newest Russian cities in 2013 were less than thirty years old, while the most ancient ones were well over a thousand. The latter include such historical cities of Ancient Rus as Veliky Novgorod, Smolensk, Pskov, Rostov, Vladimir, Bryansk, Belozersk, and Murom. What is important here is that these cities are not in ruins; quite opposite, they are living cities and relatively densely populated ones. Nevertheless, they are much poorer than many newer cities.

Examples of a new, rich city and an old, poor one are Nadym (46 years old and having an average monthly wage of 84,410 rubles) and Yelnya (863 years and 15,101 rubles). Although among the more than 1,000 cities in Russia there are old, rich ones, such as Moscow, and new, poor ones, there is a strong negative correlation between city age and wages.

This tendency can be illustrated in maps of modern Russia. Fig. 1 presents two maps in which the respective proportional symbols denote the ages (A) and nominal wages (B). It is readily seen that poorest cities are the oldest ones, and vice versa. All regions of European Russia, excluding the two capital cities with their neighboring regions and the extreme northern regions, are among the poorest. It is true also for Kaliningrad Oblast in the extreme western part of Russia and for Dagestan in the south. At the same time, the richest regions are those in Siberia and the Far East with cities founded in the late Soviet and even post-Soviet periods.

Fig. 1. Russian cities by their age and nominal wages in 2013

Note: The dots are proportional to the age in years (map A) and nominal wage in rubles (map B) of the cities. The sources are Ekonomika gorodov Rossii (2016), SSSR (1987), and Lappo (1998).

The map makes one think that Russia is not a typical country from the standpoint of spatial economics. Older and poorer cities tend to be bigger, which may suggest that there is no urban wage premium in Russia. In addition, a negative age-wage relationship is obviously not a universal phenomenon. In the US no similar link is observed, nor, probably, do other developed countries feature this inverse age-wage relationship.[1] However, the urban wage premium is still present in Russia, which is reflected in strongly positive size-income correlations. At the same time, the age-wage relationship is observed not only in Russia. Many countries among the former Soviet republics, including such diverse economies as Ukraine, Kazakhstan, and Lithuania, share this regularity with Russia. Therefore, one can argue that Russia, like other countries, have special features and features in common with others, which makes it possible to analyze this country within the framework of spatial economics.

One of the fundamental forces behind income distribution across cities is the increasing returns. As Davis and Weinstein put it, the increasing returns theories, which stress the importance of path

[1] See the dataset for the US Census places $Cities\ Databank^{TM}$.

dependence, imply, in particular, that "an early start in one location provides that site with advantages at each succeeding stage of locational competition" (2002, p. 1270). The advantages of older locations are based on the fact that the accumulation of population is a key source of increasing returns and that it is a process that takes a long time. In other words, older locations should be richer than newer ones. In light of that implication a strongly negative link between age and wages across cities in Russia and other post-Soviet countries seems puzzling. Evidence as to the potential process underlying this correlation could be informative for the debate about spatial patterns of economic activities. The agenda here includes questions such as why economic activity is unevenly distributed across space in terms of its intensity and rewards (Combes et al., 2008); or what drives changes over time in terms of interregional income inequalities.

This paper considers the forces behind spatial income distribution across Russian cities in the context of a number of important hypotheses in urban economics. The main question of this paper is why city age is inversely linked with income in Russia. To answer this question, we explore the way city age relates to a number of urban characteristics, and how these characteristics change with time for cities of different ages.

In our discussion we rely on the spatial equilibrium approach. This was 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. In particular, higher income in a place may result from higher productivity, which attracts 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 is offset either by higher prices, or disamenities, or both, so that the resultant individual utilities should be equal across space.[2] Using these assumptions Glaeser (2008) and Glaeser and Gottlieb (2009) constructed a model which makes it possible to determine the mechanism underlying the effect of a variable on income, population and prices. This paper extends their model to explicitly include natural resources as an input. For 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,

[2] Nominal income attracts people to cities, while consumption disamenities and housing price discourage people from living there. A more specific example of the model in which housing market is used as a congestion factor, which discourages people from moving to big cities, can be found in Lee and Li (2013).

2005; Brown et al., 2008).

To determine the data-generating process underlying the age-wage relationship, we test a number of traditional hypotheses. The first group of hypotheses suggest that newer cities are more productive. We distinguish between human capital, natural resources, and productivity amenities as potential sources of productivity advantages. Newer cities may be more productive because either they attract more skilled workers or use them more efficiently, or they are relatively resource-rich locations, or, finally, they have other productivity advantages such as available nontraded capital. Another group of hypotheses is related to consumption and construction disamenities. Newer cities may either be worse places for living, or they may feature construction disadvantages. The latter should translate into higher housing prices. In both cases, to keep people in the newer cities, employers would have to pay bonus offsetting the disamenity-related disutility or expensive housing.

Our own results are in favor of both the disamenity- and productivity-related hypotheses. Newer cities are poorer places in terms of their consumption amenities, but this disadvantage is partly offset by their construction amenities, which should make their housing more available. Thus, the net indirect disutility should be modest. At the same time, newer cities have a number of productivity advantages. The latter include their resource abundance and relative exhaustible resource abundance, higher returns on skills, and production amenities. Thus, higher incomes of the newer cities follow from both their productivity advantages and their poorer living conditions.

Though age is considered a 'fundamental dimension' of cities, this has mostly been neglected in the spatial economics literature (Giesen and Suedekum, 2014, p. 193). This reference is one of the few exceptions, although it has a different focus and is devoted to the US.[3] To the best of our knowledge, our paper is the first to document and explore the negative link between age and wage on data for Russia or any other country.

Previous work on the income differences across Russian territory displayed 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

[3]This paper examines the relationships between the age and size of cities using US data. One of their results is that older American cities are generally larger than newer ones. A similar result for US cities was presented in Dobkins and Ioannides (2001). The age of a city is related to the probability of having neighboring cities. Glaeser and Kahn (2001) examined the age effect on the decentralization of cities. Michaels et al. (2012) documented a positive relationship between 1880 population density and its subsequent growth for the middle-sized locations in the US.

Vakulenko, 2012; Herzfeld, 2008; Kholodilin et al., 2009). 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.[4] Despite the obvious advantages,[5] Russian cities were less frequently used as observational units, and evidence for income differentials across them is still scarce. Mikhailova (2011) analyzed city development in Russia due to Stalinist labor camps. According to this study, cities that received more initial investment when the camp system was working were more likely to survive as populated locations. This suggests that relatively new and rich cities existing today include those originating from the camp system and those having attracted significant investment, which in turn may underlie their productivity advantages.

The structure of the paper is as follows. The next section contains a thorough examination of the age-wage relationship in Russia and in other countries, and the change in time of the relationship. Section 3 outlines the spatial equilibrium model to guide the interpretation of the regression results. Section 4 contains the estimation results and the derived interpretation of the estimates based on the spatial equilibrium conditions. Then we present the robustness checks of the conclusions, and discuss their implications. Section 5 concludes.

2. Age and wage in Russia and other post-Soviet countries

2.1. Data

The unit of observation in our analysis is the city. The definition of a city as a location with official city status is motivated by the available data for this kind of Russian locations.[6] The main body of these data come from the commercially distributed base Ekonomika gorodov Rossii (2016). The main variable of our interest is city age. The dataset actually contains two age variables, however, these are not consistent. Therefore we have constructed two additional age variables which are consistently based on the foundation year and the year of giving city status, respectively.

[4] 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.

[5] The cities are much more homogenous with respect to their size and density than the regions. At the same time, in the industrialized world cities produce the bulk of national income, and focus on income disparity dynamics across cities could allow one to abstract from the agrarian sector and analyze processes specific for the industrialized part of the economy.

[6] A detailed discussion of the definition of "city" is given in Appendix A.1.

We use mostly the age variable based on the foundation year.[7] At the same time, to check the robustness of the inverse age-wage relationship, we use all the four age variables. Summary statistics of the age variables are presented in Table A1 in Appendix A.2. More than half of the cities in Russia are of Soviet origin making them just a few decades years old. Though there are very old cities, it is these new cities that are the decisive factor in forming Russian spatial income disparities.

The dataset contains key variables for testing the hypotheses, namely average wage, labor force, fixed capital stocks, resource stocks, and extractive output of energy and non-energy resources.[8] These serve as proxies for the uses of exhaustible and renewable resources based on the fact that energy resources are mostly non-renewable.[9] In addition to the main variables, we use an extensive set of control variables available through the same dataset. Data on geographical coordinates also used as controls are taken from Bariev (2007) for 992 observations; the remaining 97 items are taken from internet maps.

A number of important variables are available only at the regional level, in which cases the urban characteristics are proxied by the respective regional data.[10] To construct real wages, we use the regional price of the consumer basket. The average temperature in January is used as a control variable. For testing hypotheses we use the regional housing prices, and the prices for higher quality and typical quality housing, the share of the urban population with higher education, and wages by the level of education. These groups of variables, except the last, are only for cities within regions, meaning that in this case we use somewhat more disaggregated data compared with the remaining regional variables.

Sample statistics of these variables by age groups are given in Table A2 in Appendix A.5. These show that newer cities tend to pay higher nominal wages and — despite their consumer and housing prices also being higher — they pay higher real wages too.

For other post-Soviet countries our main task is to check if there is a similar relationship between income and city age anywhere else. For the data sources on city age, income, and population size or

[7]Our considerations here are that this variable is consistent when dealing with the start and therefore able to give a more accurate picture; the sample does not exclude the whole region, as for the start variables in Ekonomika gorodov Rossii (2016), so that it is more representative; and this refers to the actual city start, rather than the official establishment. For more details on how we constructed the age variables, see Appendix A.2.

[8]For details concerning their log transformation, see Appendix A.4.

[9]In addition, we use the resource extraction variables as proxies for the resource uses assuming that the cities extracting more resources are more resource-rich which should induce these cities to consume them more. Available data on the resource extraction and resource stocks supports this assumption. The stock-extraction correlation coefficient equals 0.83.

[10]For the sources and list of the regional data, see Appendix A.3 and Table A2 in Appendix A.5, respectively.

density by cities for these countries, see Appendix A.6.

2.2. Inverse age-wage relationship for Russia

We begin with estimating a number of cross-section regressions with results reported in Table 1. Our initial cross-section specification for 2013 (column 1) just captures the negative correlation between log age and log wage. The estimate suggests that a 1% increase in age is consistent with a 0.13% fall in the average wage. Given that the mean wage for 2013 was 25,950 rubles, at the mean point a 1% increase of city age results in a predicted fall in wage of 33.3 rubles. For bigger age differences, for example, Vysokovsk, aged 130, compared to Novomoskovsk, aged 84, is to have wage as the share of the latter city's equal to $\exp[-0.13\log(130/84)] \approx 0.95$.[11] In other words, half as much difference in age implies a 5% difference in the average wage in favor of the newer city. This basically equals to the actual relationship between their 2013 wages, namely $25,728$ rubles$/27,120$ rubles ≈ 0.95.

Table 1
City age and log wage, Russia

As seen in column 2, using real, rather than nominal, wages does not change the relationship dramatically. City age still predicts wages well after their correction for regional consumer prices. The next column contains the results after the inclusion of the regional dummies and geographical and demographic controls.[12]

The inverse age-wage relationship is robust to the inclusion of the regional dummies and controls. The age coefficient after controlling for regional dummies, and geographical and demographic characteristics remains highly significant. However, now its absolute value is less than half, meaning that the negative age-wage relationship holds both across and within regions, but the intra-regional

[11] This figure is under 0.94 when using the subsample of the small and medium-sized cities. See below.

[12] These controls may explain some variation in wage and be responsible for some part of the link. For example, latitude and longitude may be the basis for paying the "northern bonuses". The latter are special rewards of higher wages for living and working in the severe conditions of the extreme North, which are inherited from the Soviet times. In such regions wage is multiplied by the coefficient in a range of 1.15 and 2 depending on the region. At the same time, some cities might have been founded relatively recently in areas with a severe climate. Demographic controls are obvious covariates of both the age and wage. Population size and density, other things being equal, should correlate with wage (Combes et al., 2008) and be higher in older cities (Davis and Weinstein, 2002; Giesen and Suedekum, 2014). The other variables measure population dynamics and composition. The relative labor force may explain the negative link between the age and wage. In older cities there are to be more pensioners, and retired aged people often still work. Since their additional earnings are usually quite small, their proportion is to have a downward effect on the average wage.

relationship is weaker than the inter-regional one. In particular, within regions and controlling for the urban characteristics a twofold difference (which is slightly above a standard deviation difference) in age implies a 4% difference in the average wage in favor of the newer city. Given the range of the first age variable,[13] this suggests that, the urban characteristics and regional assignment being the same, the newest city should earn a more than 25% more than the oldest city.[14]

As follows from the results in column 4, the inverse age-wage relationship held more than a decade earlier, and the absolute value of the age coefficient was even higher. Now a twofold age difference, controlling for regional dummies and other characteristics, translates into a 6% negative wage difference. The remaining columns of Table 1 are for the same specifications, but with the second age variable. Again, all the age coefficients are significantly negative. Their absolute values are lower, but for the 2001 cross-section the coefficient values for both age variables are very similar. The same regressions with the use of the remaining age variables give similar results. As a whole, the inverse age-wage relationship is highly robust to changing definition of city age, included controls, and the year of the data used.

Adjusted R-squared indicates the high quality of the regressions and the non-trivial explanatory capacity of the controls. After their inclusion adjusted R-squared jumps from 0.1 to 0.8, which suggests that the bulk of wage inequalities across the cities can be ascribed to the respective inequalities across the regions and the differences in urban characteristics. The regression quality measure in column 1 is higher than in column 5, meaning that the first age variable explains more variation of log wage than the second.

2.3. Inverse age-wage relationship for other post-Soviet countries

Is the inverse relationship between city age and income an exclusively Russian anomaly? Available data on post-Soviet countries enables us to check if there are similar correlations elsewhere. We have estimated similar specifications for a number of countries given the data on city age, average income, and population size or density. The results are presented in Table 2.

[13] See summary statistics of the age variables in Table A1 in Appendix A.2.

[14] To check this relationship at the level of distinct regions, we run a number of regressions. Table A3 in Appendix B presents examples of regions with the significantly negative age-wage relationships. Some regions, e.g., Respublika Komi, Zabaykalsky krai, Orenburgskaya oblast, Vologodskaya oblast, or Kurskaya oblast show the interest coefficient under -0.2.

Table 2
Other post-Soviet countries

Similar links hold in at least 7 other countries. Among them Ukraine is the most similar to Russia in terms of the coefficient values, statistical significance, and the robustness of the inverse age-wage relationship. This relationship is highly significant both in the simple specification and with inclusion of log population and all the regional dummies. The second most similar age-wage relationship to that which holds in Russia is observed in Belarus. It is worth noting that Ukraine and Belarus are also the most urbanized among the others if measured in the numbers of cities. The remaining countries in the table display significant inverse age-income relationships depending on the specification. All of them have inverse relationships after the inclusion of log population, which is explained by the latter being positively correlated both with city age and income. Where it is possible we include all the regional dummies, and in all these cases, except 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 is more spectacular, so that the significant inverse relationship is observed only in the quadratic, rather that the logarithmic, specification.[15]

To sum up, according to the available data, the negative age-wage relationship holds for most post-Soviet countries. To some extent, this relationship can be attributed to their common Soviet legacies. At the same time, these countries differ dramatically in their territorial location, political regimes, and economies. Since the inverse age-income correlations are observed in such diverse countries, one can suppose that similar correlations exist elsewhere in the globe.

2.4. Income convergence

As follows from the interpretation of the logarithmic specification we use, the most intense negative relationship takes place for newer cities. However, it is possible that the nonlinearity in the relationship we are interested in is more than that captured by the logarithmic specification. To trace the relationship for Russia more closely, we localize the link on the scale of age using locally weighted regressions of log wage on two of the log age variables. The results for the two age regressors are

[15] In the long regression for Moldova the minimum average income is consistent with city age of 336.

presented in Fig. 2. We see that the slope of the curve is much steeper for smaller ages and becomes almost horizontal, if not positively sloped, for older ages. Thus, the pattern implied by the log specification is reinforced by running simple non-parametric regressions, which show higher absolute values of the age coefficients for newer and better-off cities. In particular, the most unambiguous negative association between age and wage is observed for cities founded or given city status under Soviet rule. In addition, these results mean that the actual sample of Russian cities which displays the negative age-wage relationship is restricted by relatively new and rich cities. This subsample also features populations under 200,000.

Fig. 2. Log wage and log age

Note: Simple locally weighted regressions were estimated with the bandwidth of 0.8.

The newer a city, the faster the relative growth of its age year to year, and these results imply that the year-to-year age increase should affect newer cities more. To check how income changes depending on city age, we consider the dynamic age-wage relationship. We have estimated panel regressions of log wage on a trend variable and its interaction with log city age. The sign of the interaction term indicates the direction of relative income change. A positive sign would indicate income convergence, meaning that newer cities, though richer, have slower income growth compared with older ones and vice versa. The regressions were estimated with city and year fixed effects, for nominal and real wages, and for the two age variables. As seen in Table 3, the interaction terms are highly significant, and their positive signs indicate the convergence among cities different in terms of their ages.

Table 3
Nominal and real wages and city age in dynamics, 1991-2013

3. Spatial equilibrium for inputs and rewards

To explain the inverse age-wage relationship and the related income convergence, one has to address an important challenge. When estimating the causal effects of various forces behind the income differences one deals with the simultaneity problem. For example, population affects income and prices via a number of channels such as labor supply and the available land for construction, while high income attracts more people, and increases housing prices, which means that all the three are endogenous variables (Glaeser 2008).

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-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.

3.1. The extended Glaeser-Gottlieb model

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 labor and capital.[16] 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

[16] The potential problem with this assumption is related to the fact that under the Soviet rule the spatial allocation of labor 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).

included as an input in the respective production and construction functions.[17] Thus, the functions contain four inputs: labor, traded capital, natural resources, and nontraded capital, among which all but the last are subject to an optimal decision at the local level. Appendices C.1-C.4 contain the static version of the model.

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 (A1), (A3)-(A5) in Appendix C in logarithms for the exogenous variables and substituting the coefficients into the solution one obtains the relationships between city age and the exogenous variables:[18]

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

$$\lambda_A = b_N(1 - \epsilon - \eta) + b_W(1 - \eta) - b_P - b_R\nu, \tag{2}$$

$$\lambda_\theta = b_P\alpha - b_W, \tag{3}$$

$$\lambda_\mu = b_N + b_W - b_R. \tag{4}$$

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 labor force, wage, housing price, and natural resource use.[19] Thus, having specific values of the parameters and substituting the coefficients from the respective regressions of the endogenous variables into (1)-(4) 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.

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

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

[19] The resultant linear combination in (3) is the same as that in Glaeser (2008, p. 55), while those in (1)-(2) differ in notations and using the resource coefficient. The additional coefficient defined in (4) follows from the inclusion of the new input in the production and construction functions.

As in Glaeser and Gottlieb (2009), the same system can be considered dynamically by first differencing the equations (A1), (A3)-(A5). The respective solution of the first-differenced equation system for the exogenous variables is as follows

$$\lambda_{1+g_T} = b_{1+g_N}(1-\beta-\gamma) + b_{1+g_W}(1-\gamma) - b_{1+g_R}\delta, \tag{5}$$

$$\lambda_{1+g_A} = b_{1+g_N}(1-\epsilon-\eta) + b_{1+g_W}(1-\eta) - b_{1+g_P} - b_{1+g_R}\nu, \tag{6}$$

$$\lambda_{1+g_\theta} = b_{1+g_P}\alpha - b_{1+g_W}, \tag{7}$$

$$\lambda_{1+g_\mu} = b_{1+g_N} + b_{1+g_W} - b_{1+g_R}. \tag{8}$$

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 (5)-(8) 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.

3.2. Human capital and exhaustible resource allocation

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 renewable 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 renewable 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[20] which provide them with different consumption amenities. Finally, the construction sector produces two kinds of housing with different productivity parameters.

The extended model is derived in Appendix C.6. The solution of the system (A9)-(A12) in Appendix C.6 for the exogenous variables gives the following relationships between the coefficients of

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

the regressions of the exogenous variables on an interest variable and the regression coefficients of the endogenous variables on the interest variable

$$\lambda_\Lambda = b_\iota(1-\epsilon-\eta-\nu) + b_\omega(1-\eta-\nu) - b_\pi, \tag{9}$$

$$\lambda_\psi = b_\omega + b_\iota(1-\sigma), \tag{10}$$

$$\lambda_\vartheta = b_\pi\alpha - b_\omega, \tag{11}$$

$$\lambda_\phi = b_\rho(\varsigma - 1). \tag{12}$$

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.

3.3. 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.

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 (1)-(4) and (9)-(12), 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 labor, 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 (5)-(8) 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:

$$y_{it} = b_1 year_t + b_2 year_t \times \log(age_i) + b_3 y_{it-1} + \mathbf{byr}_t + \alpha_i + \varepsilon_{it} \quad (13)$$

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 idiosyncratic 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 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 A2 in Appendix A.5). 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.[21]

Panel regression estimates to establish the comparative dynamics of these variables across different

[21] 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.

Table 4
City age and the endogenous variables

Table 5
City age and change of the endogenous variables

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.[22] 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.

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 labor 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.

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

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 A4 in Appendix D.2). The remaining parameters are borrowed from the literature. The parameter values used in the calculations are in Table 6.[23]

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 (1)-(4) 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.[24]

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

[23]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.

[24]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".

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 favor 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 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 favor of newer cities.

The estimation results and their theory-based interpretation are in favor 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.[25] Recall that the coefficients measure the point elasticicities of the respective exogenous variables with respect to city age. By the 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.[26] Most other

[25]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).

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

Table 6
The parameters

Table 7
City age and the endogenous variables

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 A5 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 stocks and the share of exhaustible resources in resource stocks are much stronger.[27] 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

[27] 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.

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 theater availability.[28] These are measured by the relative discharge from stationary sources and theater attendance per capita, respectively. A warmer climate, lower relative discharge, and higher theater 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.

The model suggests (e.g., (A6) in Appendix C.4) 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

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

variables, respectively. The results are in Table 8. Temperature in January and theater 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 theater 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 theater attendances, while their resource extraction measure should be higher. According to the sample statistics (Table A2 in Appendix A.5), newer cities are located to the north and the east, and suffer from a colder climate; other proxies for their consumption amenities such as theater 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 theater 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 (13). 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 theater attendance were more favorable 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 favorable 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 favorable 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 8
Exogenous and endogenous variables, 2013

Table 9
Age and consumption amenities, 2013

Table 10
City age and consumer disamenities, 2005-2013

4.6. Discussion

The established regularities for city age, wages, and other urban characteristics highlight the role of resource extraction and the various amenities in the income distribution among Russian cities. The most important force is resource depletion. New cities follow still abundant resource deposits, and their depletion makes these cities poorer as time goes by. As already existing cities are getting relatively poorer due to their resource depletion, new cities are founded at sites of new resources which provide these cities with temporary economic advantages.

In a sense, this pattern can be treated as a special case of the Schumpeterian creative destruction when applied to spatial input allocation and its rewards (Schumpeter, 1942; Hounshell, 1984; Florida, 1996). As notoriously exemplified by Detroit, once-rich locations can decline as a result of the falling relative importance of some products, technologies, and organizational forms, or the rise of new alternatives (Glaeser, 2011). The decline of some locations due to spatial creative destruction may lead to the rise of other locations. If the economic prosperity of some territories results from their resource abundance, a destructive move of inputs can be driven by new resources and resource deposits. A logical sequence would be: new technologies – new resources – new deposits – new rich regions.

In the short run, resource deposit depletion in some territories and their development in others can also lead to a change in spatial income distribution. In the long run, technological progress leads to a change in the relative importance of various resources and thereby the relative economic value of the respective territories. Literature on Russian urban history (Rodgers, 1974; Lappo, 1998; 2006; Markevich and Mikhailova, 2013) contains numerous examples of cities, whose foundation and/or development was motivated by their natural resource deposits. In such cases the new cities were founded in resource-rich, but often remote territories. These features made both their wages and prices increase. Subsequently, many such cities declined and in some cases became deserted because of the depletion or devaluation of their resources. Historical examples include cities founded near rich salt deposits such as Usolye-Sibirskoye, Solvychegodsk, Solikamsk, Sol-Iletsk, which, though some of them still contain abundant salt stocks, no longer benefit from them because of the changed conditions of salt trade. More recent history gives examples from extractive industries. According to Ekonomika gorodov Rossii (2016), many coal-mining locations in the regions of Kuzbass and Donbass used to be relatively well-off due to their flourishing coal-mining enterprises, but now their collieries are closed, which goes hand in hand with their diminishing relative wage and declining population. The examples of such locations are Anzhero-Sunzhensk, Gukovo, Prokopievsk,[29] Osinniki, Donetsk, Zverevo, Novoshakhtinsk, and Kizel. The latter is an example of dying town due to its closed collieries. Outside

[29]These cities after closing their collieries have been included in the special group of cities which are experiencing severe economic difficulties (cf. the government executive order No. 1398-r, 29.07.2014, "On the adoption of the list of monocities").

Russia an example of a location dying for the same reason is Ukrainsk in Ukraine. Similar tendencies are seen in the oil and gas industry in Russia. Between 2005 and 2013 a number of oil and gas extracting cities such as Pokachi, Uray, Pyt-Yakh, Nyagan experienced declining extractive output per capita along with decreasing population and relative wages. Such locations are mostly new cities paying substantial wage premium, but their income advantages tend to diminish with time.

From a more general perspective, such a process of city creation and evolution fits the long-standing pattern of territory development in Russia. As Kluchevsky (1911) put it, moving to yet undeveloped territories has always been the most typical strategy of the Russian population and state as a response to various challenges. In medieval Russia most peasants were involved in slash-and-burn agriculture. This provided rich harvests from new-ploughed soil but a rapid decline thereafter. When the harvests became low enough, the peasant community moved to another place. Naturally, this pattern of using territory was possible, given the large empty spaces. Russians were also induced by the search for other resources. They pushed northward and eastward motivated by their search for fur. When arriving at a fur-rich territory they built up zimovie — a fortified establishment designed to enable the newcomers to impose a tribute in furs on the native population. These establishments gave birth to many Russian cities (see, e.g., Kluchevsky 1911; Lappo 1998; 2006).[30]

This pattern is inherent for Russia, and, maybe, for other post-Soviet countries and some of the former planned economies and developing economies. But this is unlikely to be at work in developed countries. In particular, as mentioned, the US does not display any negative age-wage relationship. Concerning the resource-income relationship, the results in James and Aadland (2011) supported the resource curse on the county level. The resource curse resulted from a kind of crowding effect, meaning that the extractive industries were developed at the expense of manufacturing, which would have been more conducive to growth via increasing returns. At the same time, similar empirical studies on Russian data indicate that Russia benefits more than suffers from resources. Alexeev and Conrad (2009) point out that per capita income in Russia would be lower if Russia lacked extractive industries. Over the 2000s rapid growth induced by booming extractive output was not accompanied by the

[30] Apart from these incentives, people strived for more freedom or, as was the case with numerous religious enthusiasts, for solitude. In all the cases the state followed its subjects, resulting in the expansion of controlled territories or developing ones already controlled. Over the Soviet era the state used the vast space to retreat during the war or to deal with economic difficulties in peace time.

crowding effect, and, in fact, over the same period manufacturing production did increase (Dobrynskaya and Turkisch, 2009). Thus, while in Russia, due to its preceding development, the age predicts the relative resource scarcity and ensuing poor income, nothing of this kind is known for the US. These facts can explain the negative age-wage relationship for Russia and the lack of such a relationship for the US.

In the US, city age is a covariate of city size and other agglomeration measures. Giesen and Suedekum (2014) established the positive age-size correlation using a simple log specification with state fixed effects. For Russia an association between log age and log size is significantly positive when including regional dummies, or log wage, or log resource extraction as controls. As mentioned, the positive age-size correlations are observed also when using alternative specifications or an alternative measure of city size. At the same time, city size is a strong predictor of wage. Thus, like in the US, the older cities in Russia tend to be bigger and have urban wage premium, but this advantage is more than offset by their resource- and amenity-related features, more important for relative income.

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 labor 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.

Acknowledgments

The author is grateful for valuable comments to Sergey Izmalkov, Sergey Kokovin, Alexey Korolev, Grigory Kosenok, Jens Suedekum, Jacques François Thisse, Dao-Zhi Zeng, Ives Zenou. Special thanks go to the Co-Editor of this journal, Edward Glaeser for the important suggestions based on which the author has deeply revised the paper. All the views and errors are of the author alone.

Appendix A. Data

A.1. Definition of a city

There are numerous definitions of a city. For the most part, they refer to either some threshold values of population size or density, or to an administrative status (Rozenfeld et al., 2003; Henderson, 2005). The Encyclopedia of Russian cities (Lappo, 1998, p. 5) gives the official modern definition of a city in Russia as a location with population above 12,000 and the share of non-agricultural population in labor force no less than 85%. These are criteria from which the modern Russian state starts when giving city status. However, there are a lot of locations that do not meet these requirements still having city status. In 1989 among 1037 cities, 160 cities had less than 12,000 people, and there were 237 locations with more than 12,000 that did not have city status. As of 2011 in our dataset (Ekonomika gorodov Rossii, 2016) among 1056 cities with data on their population, 215 had population size less than 12,000. By 2013 there were 226.

Some cities have become under-populated as a result of the population decrease which a number of Russian regions faced during the later Soviet and early post-Soviet periods. However, most of these cities saved their city status due to their economic and/or administrative functions or other considerations on the part of the government. Some locations even got city status while not meeting the population size requirements. As a rule this is explained by their urban functions. The most exciting example of such a new born city is Magas that was given city status immediately after its foundation when only 100 people resided there, because Magas was designed to be the capital city of Ingushetiya, and in minds of the decision makers its administrative functions outweighed its population size when it came to giving it city status.

A.2. City age

The problem with the dataset from Ekonomika gorodov Rossii (2016) is that it mixes incomparable values under the same age variable. This variable has values that refer to a foundation year in some cases, and to a getting city status, or even later dates in other cases.

To determine the actual age definition that underlay the age variable in the dataset, we use the additional sources for city age. One of them was SSSR (1987). It is the latest official edition with information about all Soviet, including Russian, locations with their foundation and/or city status years. Another source is Lappo (1998) which includes a brief history of almost all the Russian cities that are in our main dataset (Ekonomika gorodov Rossii, 2016). Finally, there are 14 cities for which these three sources lack foundation or city status years, for which we use dates from Wikipedia. Using these we have constructed age variables that are consistent when dealing with the start and therefore give a more accurate picture.

Using Lappo (2006) to establish foundation years we rely on events that gave rise to the development of the location as an urban, rather than a rural, unit. These include the building of a military establishment such as a fortress; the emergence of an economic establishment related to a typically urban specialization, a settlement with urban specialization, a transport center, an administrative center, a custom, a health resort. In the case of ancient cities, like Moscow, the first mention (direct or indirect) in the historical records is treated as a city start on the consideration that for this kind of cities these are conventionally accepted dates of their origin.

As for getting city status, it is straightforward when relatively new cities are officially given their status, but it is not so clear for ancient cities. For the most part, in the dataset their foundation dates are treated as those of giving them city status on the consideration that they were never known as non-urban locations. Following this practice for ancient cities for which the year of being given city status is unknown we use their foundation years. There are 120 such cities. Some of these were founded as cities, having city status from the outset. Others have the same dates due to their great age. Famous examples of both types are Saint-Petersburg and Moscow, respectively. The exclusion is made for three Russian cities on the Black sea, Gelendzhik, Anapa, and Novorossiysk, which were founded as colonies of ancient Greek city-states. These cities existed under various regimes, and were at times not cities, so that their resuming under Russian empire is treated as the new beginning.

The sample size of cities with available values of constructed age variables is 1075. This is 22 age observations more than the main dataset, in which the values for city start are missing for one of the regions. As a result there are four variables referring to the city beginning – two from the dataset on Russian cities and two constructed from additional sources. Our age variables are generated as

2013-$founded$, where the first term is the latest year for which the data is available, and $founded$ is a date when a city was founded or was given city status.

Table A1
Summary statistics of the age variables, 2013

Summary statistics for the four age variables are presented in Table A1. The statistics by variables relate to each other in a predictive way. The first age variable has the highest mean, min and max values, and the second has the lowest. The most ancient cities by their foundation year are those of Greek origin which makes their age well over two thousand. The max value 1444 of the second age variable is for the former capital city of Dagestan, Derbent, which was a city from the beginning of the Middle Ages. Unlike the other age variables, for the first age variable the variation coefficient does not exceed unity. This is because this variable is not so dispersed among young and very old ages as the other age variables. The median values of all but the first age variables indicate that more than half of existing Russian cities were founded and/or given city status after initiation of the Soviet industrialization under Stalin in 1930.

A.3. Regional variables

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).

A.4. Construction of log output variables

Not all the cities extract any resources or produce manufactured output. Therefore, to avoid missing values when using the logarithmic specification, we take log(var + 1) where var is an output measure. In this way we take the logarithms of the extractive variables, manufacturing output, and the electricity, gas, and water output. Given that these output measures typically take 8-digit values, we can neglect the changes in relative values of the constructed variables.

A.5. Sample statistics

Table A2
Sample statistics for three age groups

A.6. Sources for the other post-Soviet countries

The data on city age in post-Soviet countries is borrowed from Wikipedia and checked using SSSR (1987), which contains all the locations of the USSR.

The data on income, and population size or density by cities is borrowed from the websites of regional branches of the national statistical agencies of the respective countries where they publish their municipal statistics. In particular, the websites contain data on the important cities of the respective region and its subregions. Thus, for the important cities we have ready data, while for small towns not in the municipal datasets we use the respective subregion data for which they serve as their administrative centers. We have found city level statistics for Ukraine, Belarus, Kazakhstan, and Tadzhikistan, economic region level statistics for Azerbaidzhan and Karabakh, district level statistics for Moldova, and county level statistics for Lithuania. Economic regions of Azerbaidzhan and Karabakh, and districts of Moldova are numerous enough compared with the number of their cities for their statistics to approximate well their city level statistics.

Ukraine

The list of websites of statistical agencies of all the regions, except Crimea and city of Sevastopol, is available at <http://www.ukrstat.gov.ua>. They contain the express releases of the regional statistical agencies for 2013. For Crimea we use the obsolete version of the website of the Crimean statistical agency at <http://gosstat.crimea.ru/2014/exp248.pdf> (accessed October, 2015). For city of Sevastopol we use statistics by the regions of Ukraine at

<https://ukrstat.org/uk/operativ/operativ2013/gdn/reg_zp_m/reg_zpm13_u.htm> (accessed July, 2016).

Belarus

Below is the list of regional statistics collections:

Socialno-Ekonomicheskoye polozhenie Brestskoy oblasti, 2014. Natsionalny Statisticheskiy Komitet Respubliki Belarus, Brest, 2015.

Socialno-Ekonomicheskoye polozhenie Vitebskoy oblasti, 2014. Natsionalny Statisticheskiy Komitet Respubliki Belarus, Vitebsk, 2015.

Socialno-Ekonomicheskoye polozhenie Gomelskoy oblasti, 2014. Natsionalny Statisticheskiy Komitet Respubliki Belarus, Gomel, 2015.

Socialno-Ekonomicheskoye polozhenie Grodnenskoy oblasti, 2014. Natsionalny Statisticheskiy Komitet Respubliki Belarus, Grodno, 2015.

Socialno-Ekonomicheskoye polozhenie Minskoy oblasti, 2014. Natsionalny Statisticheskiy Komitet Respubliki Belarus, Minsk, 2015.

Socialno-Ekonomicheskoye polozhenie Mogilevskoy oblasti, 2014. Natsionalny Statisticheskiy Komitet Respubliki Belarus, Mogilev, 2015.

Socialno-Ekonomicheskoye polozhenie g. Minsk, 2014. Natsionalny Statisticheskiy Komitet Respubliki Belarus, Minsk, 2015.

Lithuania

Average disposable income in cash and kind per month by county are available at

<http://osp.stat.gov.lt/en/statistiniu-rodikliu-analize?portletFormName=visualization&hash=d5c99a11-ea85-4da1-b0ae-e1e1f92fa34c> (accessed November, 2015)

Kazakhstan

The list of websites of regional statistical agencies is available at

<http://www.stat.gov.kz/faces/NavAbout/aboutAboutRegions?_adf.ctrl-state=10zg1csao9_37&_afrLoop=3

where one can download regional statistics collections. Below is their list:

Chislennost i zarabotnaya plata rabotnikov po vidam ekonomicheskoy deyatelnosti po Akmolinskoy oblasti, 2013. Kazakhstan Respublikasy Statistika agenttigi. Akmola oblysynyn Statistika departamenti, Akmola, 2013.

Oplata truda v Aktubinskoy oblasti 2010-2014. Statisticheskiy sbornik. Kazakhstan Respublikasy Yltyk ekonomika ministrligi Statistika komiteti. Aktebe oblysynyn Statistika departamenti, Aktebe, 2015.

Osnovnye pokazateli po trudu v Almatinskoy oblasti, 2013. Kazakhstan Respublikasy Statistika agenttigi. Akmola oblysynyn Statistika departamenti, Almaty, 2013.

Chislennost i zarabotnaya plata rabotnikov v Atyrauskoy oblasti, 2013. Atyrau oblysynyn Statistika departamenti, Atyrau, 2013.

Chislennost i zarabotnaya plata rabotnikov v razreze rayonov. 2013. Kazakhstan Respublikasy Statistika agenttigi. Shygys Kazakhstan oblysynyn Statistika departamenti, 2013.

Chislennost i zarabotnaya plata rabotnikov po formam sobstvennosti i vidam ekonomicheskoy deyatelnosti i v razreze rayonov. 2013. Kazakhstan Respublikasy Statistika agenttigi. Zhambyl

oblysynyn Statistika departamenti, 2013.

Osnovnye pokazateli po trudu. Kazakhstan Respublikasy Statistika agenttigi. Batys Kazakhstan oblysynyn Statistika departamenti, 2013.

Chislennost i zarabotnaya plata rabotnikov v Karagandinskoy oblasti. 2013. Kazakhstan Respublikasy Statistika agenttigi. Karagandy oblysynyn Statistika departamenti, 2013.

Chislennost i zarabotnaya plata rabotnikov po vidam ekonomicheskoy deyatelnosti Kostanayskoy oblasti. 2013. Kazakhstan Respublikasy Statistika agenttigi. Kostanay oblysynyn Statistika departamenti, 2013.

Chislennost i zarabotnaya plata rabotnikov v razreze rayonov po Kyzylordinskoy oblasti. 2013. Kazakhstan Respublikasy Statistika agenttigi. Kyzylorda oblysynyn Statistika departamenti, 2013.

Trud i zanyatost naseleniya Mangistauskoy oblasti 2009-2013. Statisticheskiy sbornik. Kazakhstan Respublikasy Statistika agenttigi. Mangystau oblysynyn Statistika departamenti, Aktau, 2013.

Chislennost i zarabotnaya plata rabotnikov Pavlodarskoy oblasti v razreze rayonov. 2013. Kazakhstan Respublikasy Statistika agenttigi. Pavlodar oblysynyn Statistika departamenti, 2013.

Regiony Severo-Kazakhstanskoy oblasti 2013. Statisticheskiy sbornik. Soltystik Kazakhstan oblysynyn Statistika departamenti, Petropavl 2014.

Yuzhno-Kazakhstanskaya oblast i ee regiony 2013. Departament statistiki Yuzhno-Kazakhstanskaya oblasti.

Tadzhikistan

Regiony Respubliki Tadzhikistan 2014. Agentstvo po statistike pri Prezidente Respubliki Tadzhikistan.

<http://stat.tj/ru/img/c476d0ddf3ea68cc502ba419312173a6_1418985745.pdf> (accessed October, 2015)

Azerbaidzhan and Karabakh

Average monthly nominal wages and salaries of employees by economic regions are available at

<http://www.stat.gov.az/source/labour/en/009_2en.xls> (accessed October, 2015)

Moldova

Principalii indicatori social-economici pe regiuni de dezvoltare, raioane si municipii. In Ianuarie-Septembrie 2014 (informatie operativa trimestriala). Biroul National de Statistica Al Republicii Modlova, Chisinau.

<http://www.statistica.md/public/files/publicatii_electronice/Principalii_ind.soc_econ/2014.zip> (accessed October, 2015)

Appendix B. Age-wage relationship by regions

Table A3
City age and real wage, distinct regions

Appendix C. The spatial equilibrium model

C.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:

$$\bar{U} = \xi \theta W P^{-\alpha}. \tag{A1}$$

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 (A1) 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:

$$D = \frac{\alpha W N}{P}. \tag{A2}$$

This will be used when solving for equilibrium housing output.

C.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 labor, 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 $TZ^{1-\beta-\gamma-\delta}$.

Labor 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.[31] We assume that, unlike labor 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 $T N^{\beta} K^{\gamma} R^{\delta} Z^{1-\beta-\gamma-\delta} - WN - K - R\mu$.

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

$$W = k_1 \left(T N^{-1+\gamma+\beta} R^{\delta} Z^{1-\beta-\gamma-\delta} \right)^{\frac{1}{1-\gamma}}, \tag{A3}$$

[31]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.

$$R = k_2 \left(TN^{\beta} Z^{1-\beta-\gamma-\delta} \mu^{-1+\gamma}\right)^{\frac{1}{1-\gamma-\delta}} . \tag{A4}$$

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.

C.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,[32] 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 (A2) one has the equilibrium housing price

$$P = k_3 \frac{N^{1-\epsilon-\eta-\nu}W^{1-\eta-\nu}\mu^{\nu}}{AL^{1-\epsilon-\eta-\nu}} \tag{A5}$$

where constant $k_3 = \alpha^{1-\epsilon-\eta-\nu}$. As a whole, we have the two linearly dependent equilibria for goods and housing, the latter being explicitly given in (A5).

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

C.4. The effect of city age on the exogenous variables

The equations (A1), (A3)-(A5) 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 (A1), (A3)-(A5) and solving it for the endogenous variables we obtain linear equations with the coefficient matrix as follows:

$$\mathbf{c} = \frac{1}{k_4}\begin{pmatrix} 1+\alpha(\eta+\nu-1) & \alpha(1-\gamma-\delta) & 1-\gamma-\delta & \delta(\alpha(1-\eta)-1)+\nu(\alpha(\gamma-1)) \\ \alpha(1-\epsilon-\eta-\nu) & \alpha(\beta+\gamma+\delta-1) & \beta+\gamma+\delta-1 & \alpha(\delta(\epsilon+\eta-1)+\nu(1-\beta-\gamma)) \\ 1-\epsilon-\eta-\nu & \beta+\gamma+\delta-1 & \beta(1-\eta-\nu)+\epsilon(\gamma+\delta-1) & \delta(\epsilon+\eta-1)+\nu(1-\beta-\gamma) \\ 1-\alpha\epsilon & \alpha\beta & \beta & \alpha(\epsilon(1-\gamma)+\beta(\eta-1))-(1-\beta-\gamma) \end{pmatrix} \tag{A6}$$

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 $\mathbf{c}$) are to positively affect all the endogenous variables. The construction amenities (the second column of $\mathbf{c}$) 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 $\mathbf{c}$) 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 $\mathbf{c}$) makes it rational to use the resources less (c_{44}), which decreases the labor 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.

C.5. The dynamic system

The first differences of equations (A1), (A3)-(A5) are as follows

$$\frac{P_{t+1}}{P_t} = \kappa_4 \left[(1+g_\theta) \frac{W_{t+1}}{W_t} \right]^{\frac{1}{\alpha}},$$

$$\frac{W_{t+1}}{W_t} = \kappa_2 \left[\frac{(1+g_T)\left(\frac{R_{t+1}}{R_t}\right)^{\delta}}{\left(\frac{N_{t+1}}{N_t}\right)^{1-\beta-\gamma}} \right]^{\frac{1}{1-\gamma}}, \tag{A7}$$

$$\frac{R_{t+1}}{R_t} = \kappa_3 \left[\frac{(1+g_T)\left(\frac{N_{t+1}}{N_t}\right)^{\beta}}{(1+g_\mu)^{1-\gamma}} \right]^{\frac{1}{1-\gamma-\delta}},$$

$$\frac{N_{t+1}}{N_t} = \kappa_1 \left[\frac{(1+g_A)\frac{P_{t+1}}{P_t}}{(1+g_\mu)^{\nu}\left(\frac{W_{t+1}}{W_t}\right)^{1-\eta-\nu}} \right]^{\frac{1}{1-\epsilon-\eta-\nu}}$$

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.[33]

C.6. Spatial equilibrium for input structure

The consumer problem is the same as in (A1), 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.

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

The construction sector uses the technology

$$A(\psi N_S^\sigma + N_U^\sigma)^{\frac{\epsilon}{\sigma}} K^\eta (R_e^\varsigma + R_r^\varsigma)^{\frac{\nu}{\varsigma}} L^{1-\epsilon-\eta-\nu} \tag{A8}$$

where ψ is productivity parameter, σ and ς denote the substitution parameter between skilled labor N_S and unskilled labor 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 (A8), FOCs for labor, capital, and resources, and housing demand functions based on the indifference conditions, as in the derivation of (A5). 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

$$\omega = \left[\frac{\iota^{\alpha(1-\epsilon-\eta-\nu)}}{\vartheta \Lambda^\alpha} \right]^{\frac{1}{1+\alpha(\eta+\nu-1)}} \tag{A9}$$

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:

$$\omega = \psi \iota^{\sigma-1}, \tag{A10}$$

$$\rho = \phi^{\frac{1}{\varsigma-1}}. \tag{A11}$$

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:

$$\pi = \frac{\iota^{1-\epsilon-\eta-\nu} \omega^{1-\eta-\nu}}{\Lambda}. \tag{A12}$$

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

Appendix D. Production function

D.1. The data

For estimating the production function we use the data from Ekonomika gorodov Rossii (2016). The production function is estimated separately for industrial production and for construction. The industrial production variable equals the annual nominal output of the manufacturing and extractive industries, and the electricity, gas, and water industry. The construction variable equals the annual nominal output of the construction industry. The input variables are labor force, fixed capital stocks, and resource stocks. As labor force we use the number of workers actually employed. The remaining inputs are measured in rubles.

D.2. The results

Table A4
Production function estimation

Appendix E. City age and the endogenous variables by subsamples

Table A5
City age and the endogenous variables

References

[1] Alexeev, Michael, Conrad, Robert, 2009. The elusive curse of oil. Review of Economics and Statistics 91 (3), 586-598.

[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] Combes, Pierre-Philippe, Mayer, Thierry, Thisse, Jacques-Francois, 2008. Economic geography. The integration of regions and nations. Princeton University Press, Princeton.

[10] Davis, Donald R., and David, E. Weinstein (2002) Bones, bombs, and break points: the geography of economic activity. American Economic Review 92 (5), 1269-1289.

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

[12] Dobrynskaya, Victoria, Turkisch, Edouard, 2009. Is Russia sick with the dutch disease? CEPII No. 20.

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

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

[15] Florida, Richard, 1996. Regional creative destruction: production organization, globalization, and the economic transformation of the midwest. Economic Geography 72 (3), 314-334.

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

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

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

[19] Glaeser, Edward L., 2011. Triumph of the city. How our greatest invention makes us richer, smarter, greener, healthier, and happier. Penguin Press, New York.

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

[21] 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.

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

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

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

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

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

[27] Henderson, J. Vernon, 2005. Urbanization and growth. Handbook of Economic Growth 1, 1544-1591.

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

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

[30] Hounshell, David, 1984. From American system to mass production, 1800-1932. Johns Hopkins University Press, Baltimore.

[31] James, Alex, Aadland, David, 2011. The curse of natural resources: an empirical investigation of U.S. counties. Resource and Energy Economics 33 (2), 440-453.

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

[33] Kluchevsky, Vasily. O., 1911. History of Russia. Vol. I. J. M. Dent & Sons, Ltd, London.

[34] 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.

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

[36] Lappo, Georgy M., 2006. Vozrast gorodov. Geographia, 7-8.

[37] Lee, Sanghoon, Li, Qiang, 2013. Uneven landscapes and city size distributions. Journal of Urban Economics 78 (1), 19-29.

[38] 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.

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

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

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

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

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

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

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

[46] Rodgers, A., 1974. The locational dynamics of Soviet industry. Annals of the Association of American Geographers, 64 (2), 226-240.

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

[48] Rozenfeld, Hernan D., Rybski, Diego, Gabaix, Xavier, Makse, Hernan A., 2011. The area and population of cities: new insights from a different perspective on cities. American Economic Review 101 (5), 2205-2225.

[49] 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.

[50] 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).

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

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

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

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

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

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

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

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

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

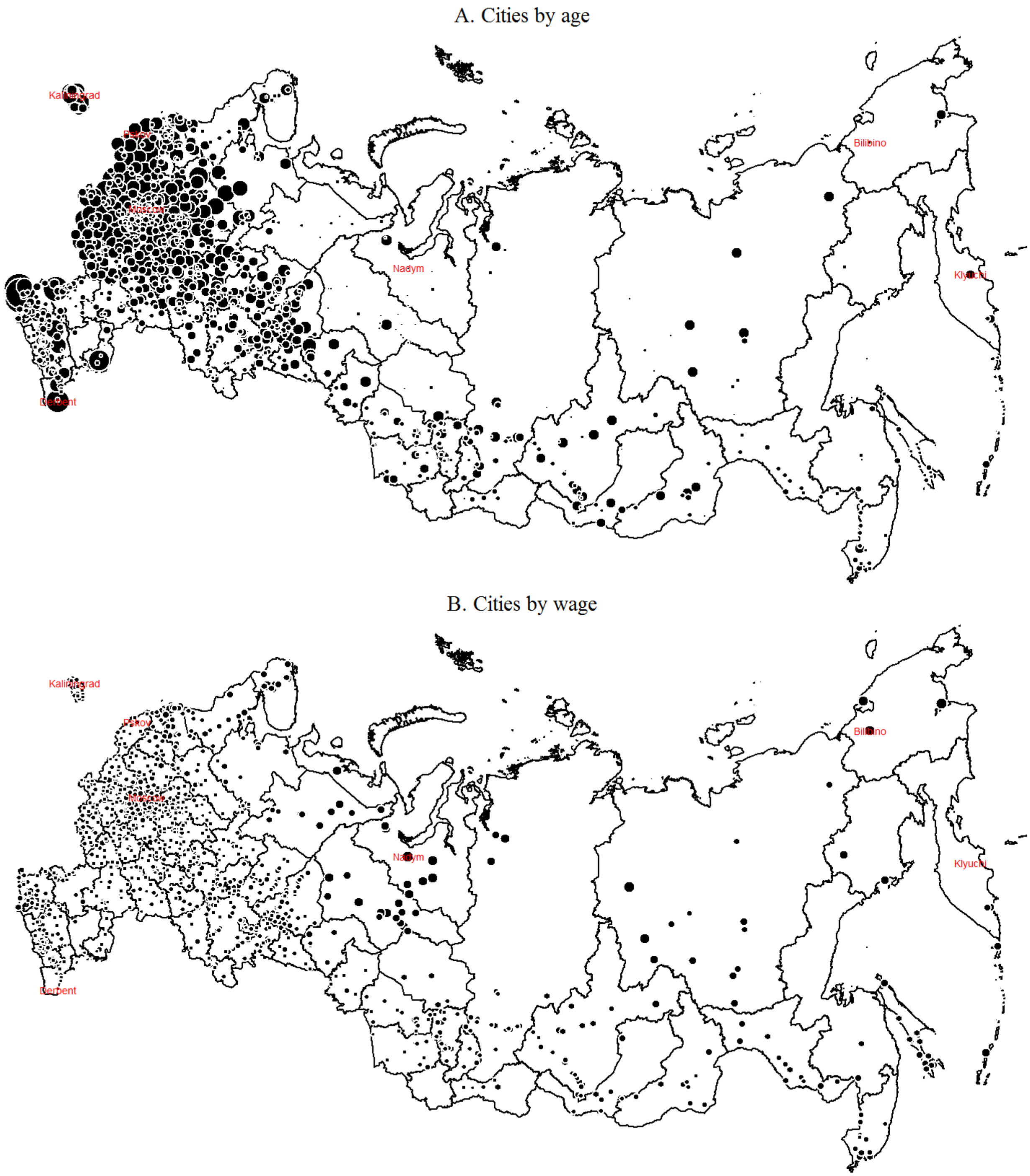

Fig. 1. Russian cities by their age and nominal wages in 2013
Note: The dots are proportional to the age in years (map A) and nominal wage in rubles (map B) of the cities. The sources are Ekonomika gorodov Rossii (2016), SSSR (1987), and Lappo (1998).

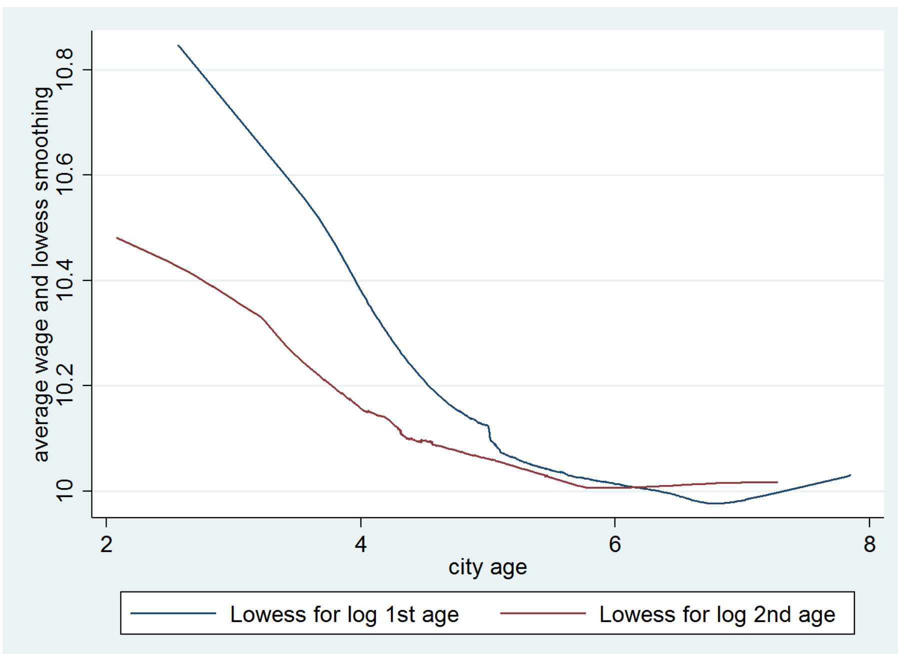

Fig. 2. Log wage and log age
Note: Simple locally weighted regressions were estimated with the bandwidth of 0.8.

Table 1
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

Table 2
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 sandwich 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.

Table 3

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 sandwich 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

Table 4

City age and the endogenous variables

Panel A	Dependent variable, 2013			
	log labor	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 labor force to commercial firms ratio. *** p<0.01, ** p<0.05, * p<0.1

Table 5

City age and change of the endogenous variables

Panel A	Dependent variable, 1996-2013			
	log labor	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

Table 6
The parameters

parameter	value	source
α	0.103	The EU in the world (2015, p. 35)
β	0.6946	Table A4 in Appendix D.2; Kuboniwa (2011, p. 8), Rõõm (2001, p. 10)
γ	0.0715	Table A4 in Appendix D.2
δ	0.1338	Table A4 in Appendix D.2
ϵ	0.7416	Table A4 in Appendix D.2; Serebryakov (2000, p. 157)
η	0.0554	Table A4 in Appendix D.2
ν	0.103	Table A4 in Appendix D.2
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 7
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 (**??**)-(**??**), and (**??**)-(**??**).
The coefficients in columns 3 and 4 for $\log(\psi)$, $\log(\Lambda)$, $\log(\vartheta)$, $\log(\phi)$ were calculated according to (**??**)-(**??**)

Table 8
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]
Theater 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 9

Age and consumption amenities, 2013

	Temperature in January	Discharge ratio	Theater 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 10

City age and consumer disamenities, 2005-2013

	Temperature in January	Discharge ratio	Theater 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

Table A1

Summary statistics of the age variables, 2013

Variables	obs	mean	median	sd	min	max
City age (since a city origin)	1,075	289.7	208	277.6	13	2,563
City age (since a city status)	1,075	182.6	77	219.9	8	1,446
City age (from the dataset with some earlier dates)	1,053	201.4	81	240.0	9	1,574
City age (from the dataset with some later dates)	1,053	196.9	78	238.3	9	1,574

Notes: the dataset is Ekonomika gorodov Rossii (2016)

Table A2
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
Labor 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									
Theater 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
Labor 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 labor ratio	996	40.3453	130.2874	56	160.2853	408.7728	42	111.3335	328.2192
Manufacturing output to labor ratio	996	209.2253	406.3291	56	198.6823	369.0112	42	48.9284	114.5952
Extractive output to labor ratio	996	53.4801	472.6981	56	437.5279	1,172.8821	42	179.8800	472.8771
Energy to non-energy 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 non-energy 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.

Table A3

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 sandwich 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 center, and latitude and longitude.

Table A4

Production function estimation

dependent variable	Log output	Log construction
	(1)	(2)
Log labor 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 sandwich 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.

Table A5

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.