Journal of Urban Economics 104 (2018) 16–34

Contents lists available at ScienceDirect

Journal of Urban Economics

journal homepage: www.elsevier.com/locate/jue

Why do newer cities promise higher wages in Russia?

Alexander S. Skorobogatov*

National Research University Higher School of Economics, St. Petersburg, Russia

A B S T R A C T

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 age-related urban characteristics responsible for this relationship, we develop a spatial equilibrium model as a framework to guide the interpretation of the regression estimates. 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 amenities, higher shares of skilled workers, and more available natural resources. These advantages and disadvantages tend to disappear over time, which gives rise to income convergence.

1. Introduction

Russian territories display an 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.[1] Examples of a new, rich city and an old, poor one are Nadym (46 years old and having an average monthly wage of $2,574) and Yelnya (863 years and $460).[2] Although among the more than 1000 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 Moscow and St. Petersburg and their neighboring regions, and the extreme northern regions, are among the poorest. It is true also for Kaliningrad region in the extreme western part of Russia and for Dagestan in the south. The richest regions are those in Siberia and the Far East with cities founded in the late Soviet and even post Soviet periods.

Russia does not seem to be 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. This negative age-wage relationship is not a universal phenomenon. In the US no similar link is observed, nor, probably, do other developed countries feature this inverse age-wage relationship.[3] However, the urban wage premium is still present in Russia, which is reflected in strongly positive size-income correlations. In addition, the age-wage relationship is observed not only in Russia. Many countries among the former Soviet republics, including such diverse economies as Ukraine and Kazakhstan share this regularity with Russia. Therefore, one can argue that Russia, like other countries, has special features, and features in common with others, which makes it possible to analyze this country within the framework of spatial economics.

The new economic geography considers increasing returns as a fundamental force behind income distribution. As Davis and Weinstein put it, this general approach includes theories which stress the importance of path dependence and imply 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 for the potential processes underlying this correlation could be informative for the debate about spatial patterns of economic activities. The agenda in the field 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.

* Corresponding author.
E-mail address: skorobogat@mail.ru.

[1] 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 the Ancient Rus as Veliky Novgorod, Smolensk, Pskov, Rostov, Vladimir, Bryansk, Belozersk, and Murom. These cities are not in ruins; quite opposite, they are living cities and relatively densely populated.

[2] Here and below we use the official RUB to USD exchange rate of 32.8 in 2013 December.

[3] See the dataset for the US Census places *Cities Databank™*.

https://doi.org/10.1016/j.jue.2017.12.003
Received 3 September 2016; Received in revised form 28 November 2017; Accepted 20 December 2017
Available online 24 December 2017
0094-1190/

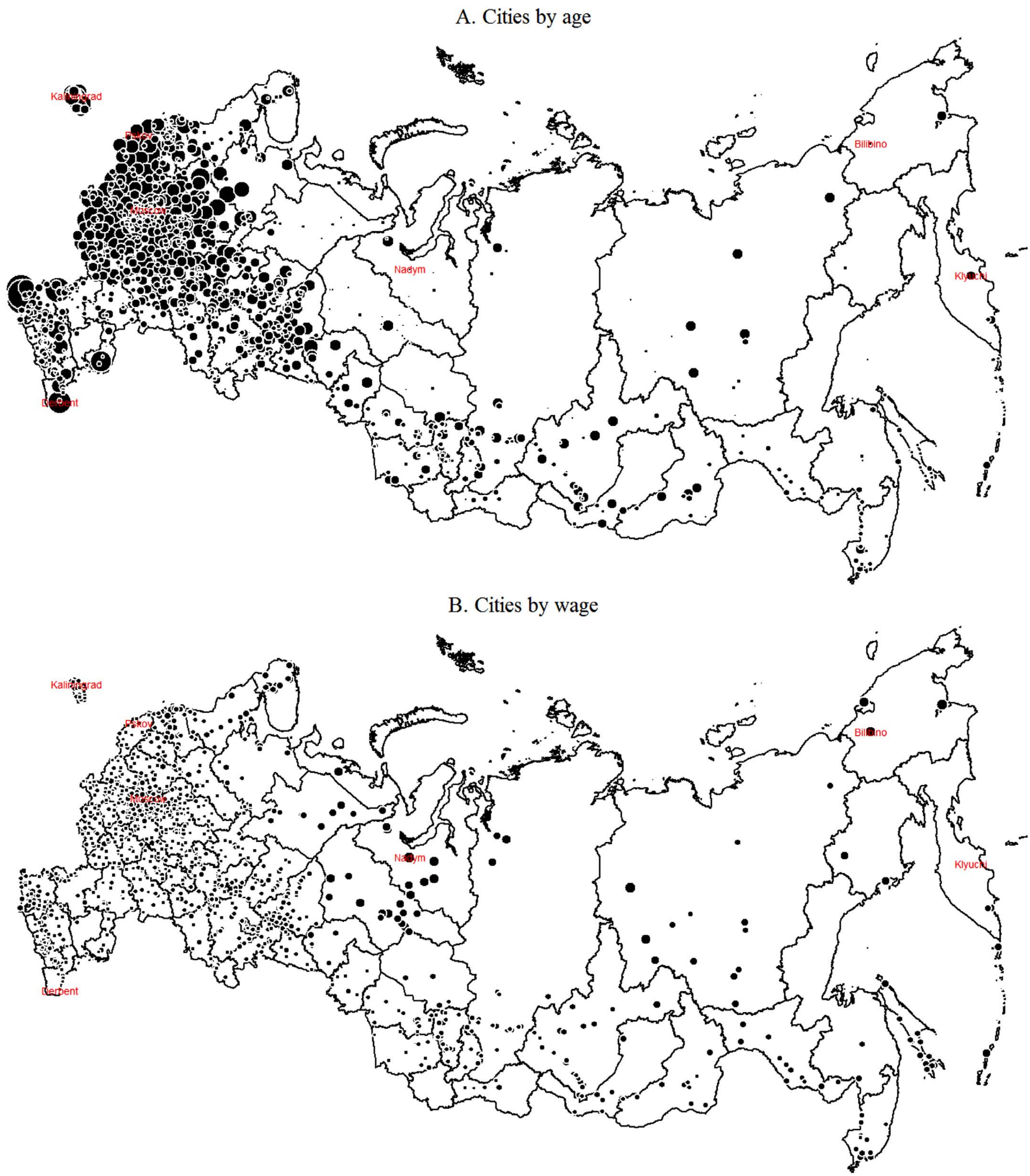

Fig. 1. Russian cities by their age and nominal wages, 2013 *Notes*: The dots are proportional to the age in years (map A) and nominal wage in rubles (map B) of the cities. The sources are Ekonomika gorodov Rossii (2016); SSSR (1987), and Lappo (1998).

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 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, which has been extensively used to deal with similar problems (e.g., Autor and Dorn, 2013; Beaudry et al., 2014; Glaeser and Gottlieb, 2009; Moretti, 2013). Its main assumption is 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, which in turn increases housing prices. Another reason for higher income may be disamenities. Ultimately, higher income is offset either by higher prices, or disamenities, or both, so that the resultant individual utilities should be equal across space.[4]

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 include natural resources as an input extracted locally and used for the production of a finished good. 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, 2005; Brown et al., 2008).

To determine the age-related urban characteristics 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 consider human capital, natural resources, and production 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 to live in, or they may feature construction disadvantages. The latter should translate into higher housing prices. In both cases, to keep people in newer cities, employers would have to pay a premium offsetting the disamenity-related disutility or expensive housing.

Our results favor 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 housing more available. On the other hand, 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. We explain these results by a kind of spatial creative destruction. Local characteristics, which made cities productive in the past, deteriorate or depreciate, while the living conditions of these cities improve over time.

Though age is considered a ‘fundamental dimension’ of cities ((Giesen and Suedekum, 2014), p. 193), this has mostly been neglected in the spatial economics literature. The work by Giesen and Suedekum (2014) is one of the few exceptions, although it has a different focus and is devoted to the US.[5] To the best of our knowledge, our paper is the first to document and explore the negative link between age and wage for any country.

Previous work on the income differences across Russia showed income convergence among Russian regions after 2000, spatial clustering in terms of regional incomes, and the important role of natural resources for the income gaps across the regions (Dolinskaya, 2002; Guriev and 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.[6] Despite the obvious advantages,[7] 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 the relatively new and rich cities existing today include those originating from the camp system and those having attracted significant investment, which may underlie their productivity advantages.

The structure of the paper is as follows. Section 2 examines the age-wage relationship in Russia and in other countries, and the change of this relationship over time. Section 3 outlines the spatial equilibrium model to guide the interpretation of the regression results. Section Section 4 contains the estimation results and the derived interpretation of the estimates based on 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. The main body of these data come from the commercially distributed base Ekonomika gorodov Rossii (2016). The variable of our interest is city age. We have constructed two age variables based on the foundation year and the year of receiving city status. The first age variable is used to establish the main correlation and to test the hypotheses, while the second age variable to distinguish the age groups and to check the robustness of the correlation of interest.[8]

The dataset contains key variables for testing the hypotheses, namely average wage, labor force, fixed capital stocks, resource stocks, manufacturing output, construction output, and extractive output of energy and non-energy resources.[9] The latter serve as proxies for the uses of exhaustible and renewable resources based on the fact that energy resources are mostly non-renewable. In addition to the main variables, we use an extensive set of control variables available through the same dataset. In particular, we use consumer amenities such as the quality of the natural and cultural environments measured by the relative emissions of air contaminants from stationary sources and theater attendance per capita, respectively. Data on geographical coordinates are also used as controls and 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 for regions is used as a control and an interest variable. For testing hypotheses we use regional housing prices, and prices for higher quality

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

[5] They examine 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 Yannis (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 subsequent growth for middle-sized locations in the US.

[6] In particular, the 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 in size to countries.

[7] Cities are much more homogenous with respect to their size and density than regions. In addition, in the industrialized world they produce the bulk of national income, and a focus on income disparity dynamics across cities would allow one to abstract from the agrarian sector and analyze processes specific to the industrialized part of the economy.

[8] We treat the first variable as the main one because this refers to the actual city start, rather than the official establishment. The second variable, being mostly a shift ahead in time, is a close covariate of the first age variable with their correlation coefficient of 0.9. The age groups are based on the second variable to have sufficient observations for the groups of new cities of Soviet origin. For a discussion of the definition of “city” and details on how we constructed the age variables, see on-line Appendices D.1 and D.2, respectively.

[9] Methodology of compiling the data on city-specific wage is discussed in on-line Appendix D.3. For details concerning log transformation of the output variables, see on-line Appendix D.6.

[10] For the sources the regional data, see on-line Appendix D.5.

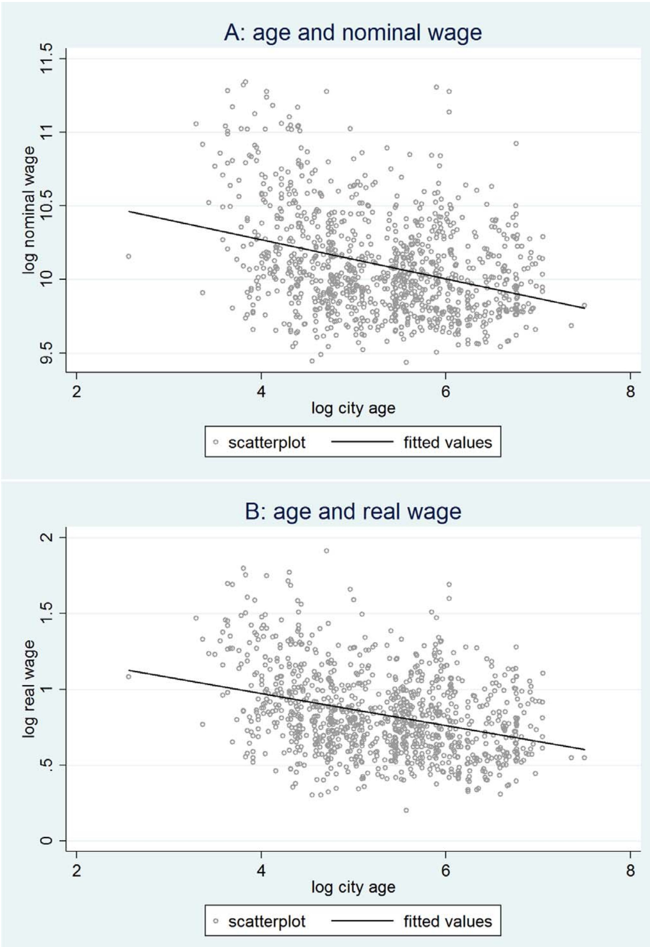

Fig. 2. City age and nominal and real wages, 2013.
Notes: The regression lines in graphs A and B are $\log Wage = 10.806[0.071] - 0.133[0.013] \times \log Age$, $N = 1046$, $R^2 = 0.11$, $\log Real\ Wage = 1.399[0.054] - 0.106[0.010] \times \log Age$, $N = 1046$, $R^2 = 0.11$. Robust standard errors obtained by the sandwich estimator are in brackets.

and typical quality housing, the share of the urban population with higher education, and wages by level of education. These groups of variables, except the last, are only for cities within regions, meaning that in these cases we use somewhat more disaggregated data compared with the remaining regional variables. Sample statistics of all the variables by age groups are given in Table A5 in on-line Appendix D.4.

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 density by cities for these countries, see on-line Appendix D.7.

2.2. *The inverse age-wage relationship for Russia*

We begin with estimating simple regressions of nominal and real wages. The estimation results and scatterplots are in Fig. 2.

The first 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 $791, at the mean point a 1% increase of city age results in a predicted fall in wage of $1.02. Given the range without outliers of the first age variable,[11] the estimate suggests that the newest city should earn a more than 40% more than the oldest city.

Scatterplot B in the same figure shows that using real, rather than nominal, wages does not change the relationship dramatically. City age still predicts wages well even after their correction for regional consumer prices.

To control for other urban characteristics, we run a number of long cross-section regressions with results reported in Table 1.

One of the characteristics that should correlate with city age is city size, the reason being that the accumulation of population takes time. In addition, people move to a city attracted by its comparative characteristics, therefore its size is a strong covariate of other important urban characteristics. Column 1 of Table 1 contains the results after the inclusion of log population size. This control is highly significant, meaning that bigger cities tend to be richer. Its inclusion substantially improves the goodness of fit. The interest coefficient remains significantly negative, its absolute value being slightly higher compared with the simple regression. To trace the relationship between city size and the age-wage correlation more closely, we run a number of simple regressions for groups of cities differing by size. The scatterplots and the estimates are presented in Fig. 3.

Cities of different sizes show an inverse age-wage relationship, though the slope is different from one group to another. The highest slopes are demonstrated by cities of different sizes. While the highest slope is seen in the group of cities with the size within the 0.1 quantile, the second highest slope is demonstrated by the cities with the size between the 0.5 and 0.75 quantiles.

Apart size, cities differ in geographical characteristics. To check how the latter affect the age-wage link, we run the respective regressions for the broad regions of European Russia and Asian Russia. The estimates and scatterplots are in Fig. 4.

While the negative age-wage relationships in both regions are highly significant, the relationship in Asian Russia is more pronounced. This suggests that geographical characteristics may be partly responsible for the age-wage link.[12] The estimation results after the inclusion of the geographical and demographic controls in column 2 confirm this suggestion.[13]

When we add the respective controls the absolute value of the interest coefficient becomes lower, though it is still highly significantly negative. The next three columns of the table present results after adding amenity-related, economic, and industrial characteristics, respectively. Though these controls improve the quality of regression, the interest coefficient does not change much. The age dummies allow us to check whether the link holds within samples of cities by age. Their inclusion slightly changes the coefficient and improves the regression quality. As an additional robustness check we also run a number of simple regressions for respective samples. The estimates and scatterplots are in Fig. 5. The interest link holds within all the age groups we use.

Finally, the results in column 7 of Table 1 show that the inverse age-wage relationship is also robust to the inclusion of the regional dummies. The age coefficient, after controlling for regional dummies and controls, remains highly significant. However, its absolute value became still lower, meaning that the negative age-wage relationship holds

[11] This range is 13–599.

[12] A positive age-size correlation is not observed for European Russia, but is for Asian Russia. After the inclusion of population size the coefficient for the latter group changes quite modestly from − 0.18 to − 0.19.

[13] 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 have been inherited from Soviet times. In such regions wages are multiplied by a coefficient of 1.15–2 depending on the region. 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 should correlate with wage (Combes et al., 2008) and be higher in older cities (Davis and David, 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 more pensioners, and retired aged people often still work. Since their additional earnings are usually quite small, their proportion has a downward effect on the average wage.

Table 1
City age and log wage in Russia, 2013.

	+ City size	+ Geo and demog	+ Amenities	+ Economic	+ Industrial	+ Age dummies	+ Regional dummies
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log age	− 0.1347***	− 0.0802***	− 0.0762***	− 0.0747***	− 0.0716***	− 0.0798***	− 0.0500***
	[0.0120]	[0.0095]	[0.0096]	[0.0098]	[0.0097]	[0.0142]	[0.0139]
Log population size	0.0857***	0.0425***	0.0475***	0.0476***	0.0325***	0.0357***	0.0432***
	[0.0076]	[0.0098]	[0.0097]	[0.0093]	[0.0110]	[0.0112]	[0.0084]
Observations	1,046	933	933	872	872	872	872
R^2 adj.	0.206	0.634	0.645	0.712	0.715	0.717	0.843

Note: Robust standard errors clustered at the region level are in brackets. *** p < 0.01, ** p < 0.05, * p < 0.1. Geographical controls include latitude and longitude, log distances to railroads and docks, and a dummy for the status of regional administrative center. Demographic controls, in addition to log population size, include log density, net migration per capita, relative male, youth, working age, and elderly populations, birth rate, mortality, and average household size. Amenity controls include temperature in January, relative emissions of air contaminants from stationary sources, and theater attendance per capita. Economic controls include log unemployment rate, log commercial firms to all organizations ratio, and log labor force to commercial firms ratio, relative numbers of college students and unemployed having completed skills programs. Industrial controls include the logs of manufacturing, construction, and extraction outputs. Age dummies are for cities with city status before 1703, between 1703 and 1917, between 1917 and 1930, between 1930 and 1953, between 1953 and 1982, and after 1982. Regional dummies are for 83 regions as of 2013.

both across and within regions. Within regions and controlling for the urban characteristics, in particular, a twice difference (which is slightly more than a standard deviation difference) in age implies a 3% difference in the average wage in favor of the newer city. This suggests that, the urban characteristics and regional assignment being the same, the newest city should earn a more than 20% more than the oldest city (ignoring outliers).[14]

The inverse age-wage relationship hold for all the available cross-sections between 1991 and 2013, and for many of them the absolute values of the age coefficient are even higher. For example, according to the estimate we omit here, in 2001 the newest city should have earned more than 53% more than the oldest city (ignoring outliers). In addition, the same regressions were estimated using the remaining age variables.[15] All the age coefficients are significantly negative. As a whole, the inverse age-wage relationship is highly robust to changing the definition of city age, included controls, and the year of the data used.

The adjusted *R*-squared indicates the high quality of the regressions and the non-trivial explanatory capacity of the controls. After their inclusion the adjusted *R*-squared jumps from 0.206 to 0.843, 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 same regressions were run with the use of a consistent sample. Based on the approach of Altonji et al. (2005) and Oster (2017) we also evaluated the remaining potential bias of the estimate of interest. For the results and their discussion, see on-line Appendix F. The sign of the true estimate is the same though their absolute value is lower than our coefficient, meaning that there are unobservable urban characteristics which at least partly explain the negative age-wage relationship. To determine these unobservables, we use a theoretical framework described below.

2.3. Income convergence

The newer the city, the faster the relative growth of its age year to year. Therefore, as follows from the interpretation of the logarithmic specification we use, the most intense negative relationship takes place for newer cities, meaning 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 run panel regressions of log wage on a trend variable and its interaction with log city age. The estimates are obtained controlling for the dependent variable lag, and city and year fixed effects for the age groups of cities. The sign of the interaction term indicates the direction of relative income change. The results are in Table 2.

The interaction terms are significant for all the age groups. Their positive signs indicate a convergence, meaning that newer cities, though richer, have slower income growth compared with older ones and vice versa. This pattern is robust to using much smaller age subsamples which is seen in that similar relative dynamics are observed within the age groups. The convergence also suggests the existence of a smooth negative age-wage relationship, rather than just age subsamples with different mean incomes.

2.4. The 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 with data on city age, average income, and population size or density. The results are presented in Table 3.

The results show that at least 7 other countries have the similar age-wage correlations. Three of these countries show negative age-wage relationships in the simple specification, and all of them have the negative age-wage relationships after the inclusion of log population. 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 the 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. Ukraine and Belarus are also the most urbanized among the other post-Soviet countries measured in the number of cities. The remaining countries in the table display significant inverse age-income relationships depending on the specification. Where 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 for Moldova the nonlinearity is more spectacular, so that the significant inverse relationship is observed only in the quadratic, rather that the logarithmic, specification.[16]

To sum up, according to the available data, the negative age-wage relationship holds for most post-Soviet countries. To some extent, this

[14] To check this relationship at the level of particular regions, we run a number of regressions. Fig. A1 in on-line Appendix E presents scatterplots of significantly negative age-wage relationships for several regions.

[15] Apart from the second age variable these include the raw age variables from the dataset. See on-line Appendix D.2.

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

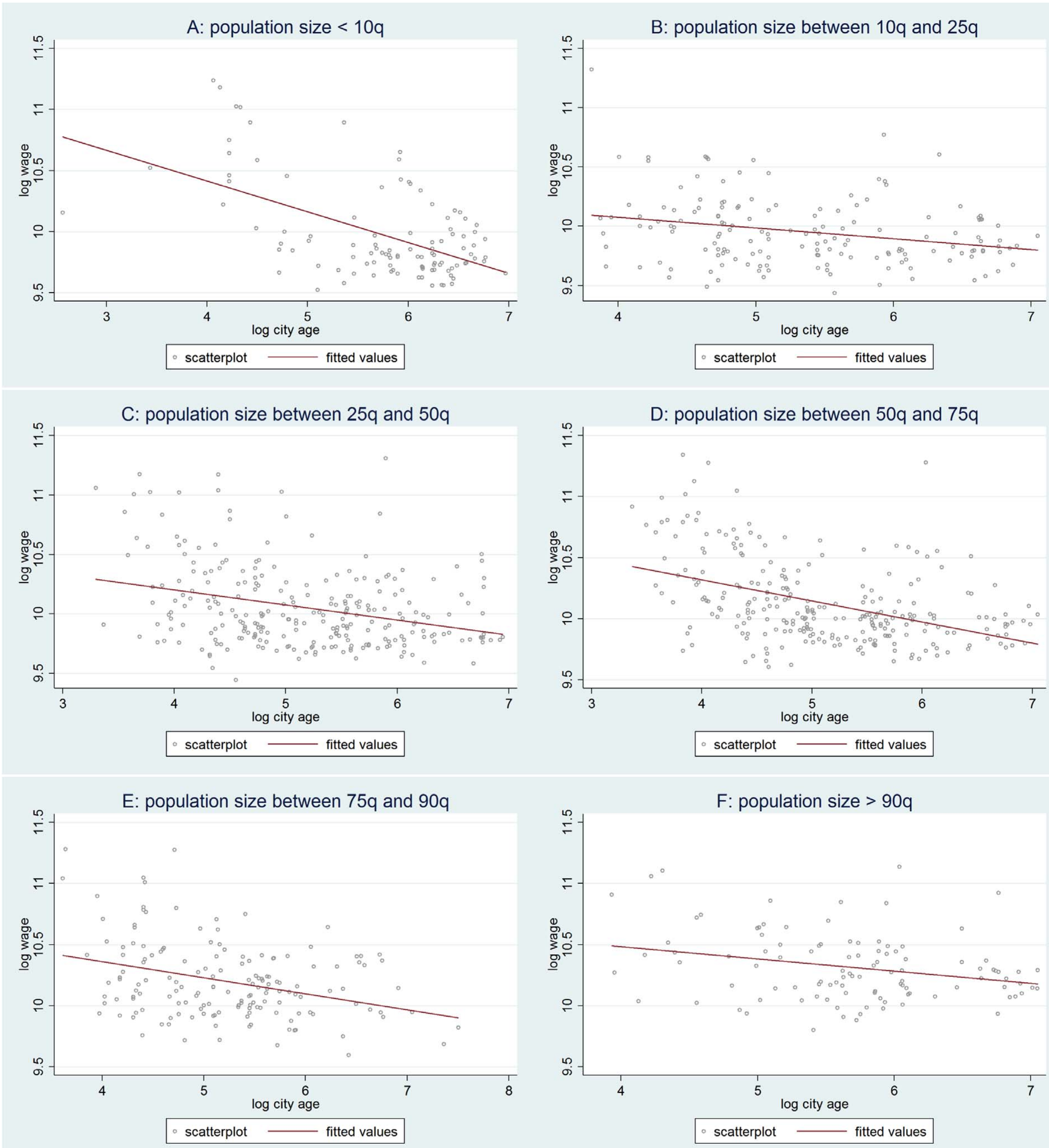

Fig. 3. Age and wage for cities of different sizes, 2013.

Notes: The samples are restricted by population size. The regression lines in graphs A-F are A: $\log Wage = 11.423[0.301] - 0.252[0.050] \times \log Age$, $N = 105$, $R^2 = 0.30$, B: $\log Wage = 10.437[0.164] - 0.090[0.029] \times \log Age$, $N = 158$, $R^2 = 0.07$ C: $\log Wage = 10.709[0.147] - 0.127[0.027] \times \log Age$, $N = 258$, $R^2 = 0.10$ D: $\log Wage = 11.011[0.127] - 0.173[0.023] \times \log Age$, $N = 261$, $R^2 = 0.20$ E: $\log Wage = 10.888[0.179] - 0.132[0.032] \times \log Age$, $N = 158$, $R^2 = 0.11$ F: $\log Wage = 10.881[0.219] - 0.099[0.037] \times \log Age$, $N = 106$, $R^2 = 0.08$ Robust standard errors obtained by the sandwich estimator are in brackets.

relationship can be attributed to their common Soviet legacy. However, these countries began their independent existence with very different backgrounds. Immediately after the collapse of the USSR in 1992 its former republics differed dramatically in their geographical location, political regimes, and economies. The subsequent histories of these countries have also been very different. By 2013 the initial deep differences had become even bigger. Since the inverse age-income correlations are observed in such diverse countries, one can suppose that similar correlations exist elsewhere in the globe.

2.5. *Does city age matter?*

We found that negative age-wage correlations hold in post-Soviet countries, including Russia. One can suppose that this results entirely from their common Soviet legacy. Newer cities are those emerged in the Soviet era for productivity-related reasons, which means that their relatively young age and high incomes simply follow from their Soviet origin and the purpose of their creation. This supposition implies a number of predictions: one would expect a sharp distinction between cities of Soviet and pre-Soviet origins in terms of average wage; a negative age-wage relationship should hold only when comparing Soviet

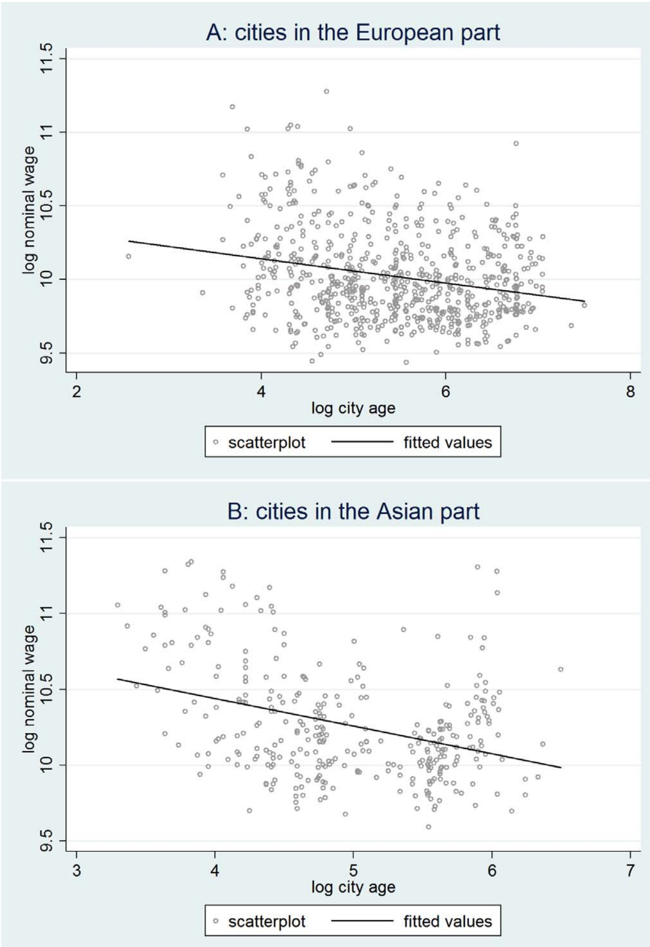

Fig. 4. Age and wage in the European and Asian parts of Russia, 2013.
Notes: The regression lines in graphs A and B are $\log Wage = 10.478[0.076] - 0.083[0.013] \times \log Age$, $N = 719$, $R^2 = 0.06$, $\log Wage = 11.168[0.152] - 0.182[0.030] \times \log Age$, $N = 327$, $R^2 = 0.13$. Robust standard errors obtained by the sandwich estimator are in brackets.

and non-Soviet samples, but not within either of them; still less probable would be a relationship of this kind within lesser subsamples of both groups; the comparative growth of average wages and urban characteristics affecting wages, such as resource abundance, should not depend on city age.

Fig. A2 in on-line Appendix shows two density plots for cities of Soviet and pre-Soviet origins. The most frequent, min and max values of the wage variable for both samples are similar. The main feature of the Soviet sample is that the average wage is observed less frequently and wages between \$725 and \$925, and higher than \$1050 are observed more often.

According to the results presented in column 6 of Table 1 and Fig. 5, the correlation holds within the samples. The negative age-wage correlation estimated using the whole population of cities is robust to controlling for whether the city is of Soviet origin and even for much narrower age subgroups. In addition, a similar age-wage correlation is observed within a number of age subsamples from Soviet and pre-Soviet samples. Cities from either sample differ from each other in their wage depending on their relative age and this correlation holds within the smaller subsamples.

Estimates of the relative income growth presented in Table 2 show an income convergence among the age subsamples, which means that the same cities change their relative income depending on their age. Between 1991 and 2013 an additional year contributed more to the relative age of newer cities, which correlated with decreasing their relative wages. In other words, there is a smooth relationship between age and wage.

Though the differences between the Soviet and non-Soviet samples make some contribution to the negative age-wage relationship, this relationship is far from being driven only by these differences. The Soviet legacy of artificial urban development plans, though important, does not fully explain this relationship. The Soviet cities created later tend to be richer, meaning that age matters for income distribution even among relatively new cities of Soviet origin. In addition, the age-wage relationship is observed not only among cities of Soviet origin, but also among older age groups. This may be explained by the fact that in the Imperial period and earlier in Russian history cities were created discretionally for some productive reasons (Lappo, 1998; 2006; Markevich and Mikhailova, 2013).

Age may affect the city's characteristics, because the latter are related to the time of city origin. Geographic characteristics may change their value as, for example, resource stocks may be exhausted, or demand for them may change with time. In addition, other characteristics may improve or deteriorate depending on human activities. Time is a force affecting a lot of important local characteristics, and this may underlie the correlations between age and wage or other characteristics.

Older cities tend to be bigger because they have more time to accumulate population.[17] Newer cities should have size-related disadvantages, which make their earnings lower. Without these disadvantages they would earn still more, meaning that they have age-related characteristics which more than offset their size-related disadvantages. Therefore, for city of equal size, the age-wage link is more pronounced. This is seen in the results for Russia and much more in those post-Soviet countries for which negative age-wage relationships hold only controlling for city size.

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). Another problem is the lack of proper data on the exogenous variables.

One of the ways to determine the mechanism underlying the effect of a variable on endogenous variables is to use a theoretical framework. The theory can suggest a relationship between the vectors of exogenous and endogenous variables $\phi: x^k \to y^k$, while the available data can give 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 variable of interest on the endogenous variables into its partial effects on the exogenous variables $\boldsymbol{\lambda}$.

3.1. The extended Glaeser–Gottlieb model

A 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 the related literature, we assume that spatial patterns of input allocation and their rewards are governed by the no-

[17] Population growth in a location depends on its changing characteristics, so that a newer location may be bigger than an older one because of more favorable changes in its characteristics in the past. In addition, since local population growth normally slows with time, age-size correlations tend to be stronger for younger cities. This explains the moderate positive correlation of 0.1 between city size and city age for Russia.

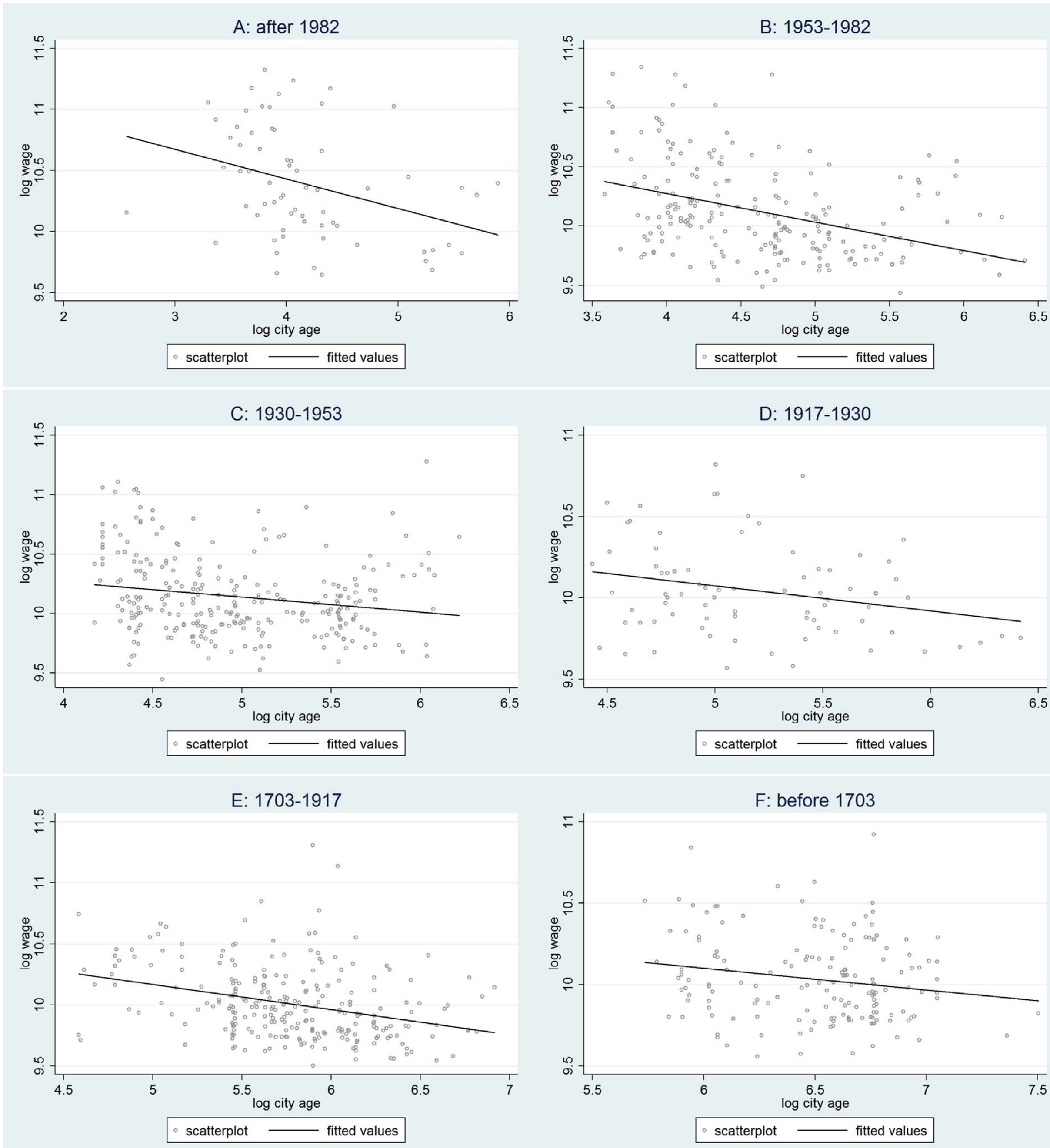

Fig. 5. Age and wage for cities of different ages, 2013.

Notes: The samples are restricted by the periods of getting city status. The regression lines in graphs A-F are A: $\log Wage = 11.400[0.331] - 0.242[0.077] \times \log Age$, $N = 66$, $R^2 = 0.12$, B: $\log Wage = 11.236[0.209] - 0.240[0.044] \times \log Age$, $N = 207$, $R^2 = 0.15$ C: $\log Wage = 10.770[0.227] - 0.126[0.045] \times \log Age$, $N = 278$, $R^2 = 0.04$ D: $\log Wage = 10.837[0.300] - 0.153[0.056] \times \log Age$, $N = 79$, $R^2 = 0.07$ E: $\log Wage = 11.200[0.217] - 0.210[0.037] \times \log Age$, $N = 252$, $R^2 = 0.11$ F: $\log Wage = 10.900[0.361] - 0.133[0.055] \times \log Age$, $N = 164$, $R^2 = 0.03$. Robust standard errors obtained by the sandwich estimator are in brackets.

arbitrage condition for consumers, firms, and developers resulting from the free movement of labor and capital.[18] 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. As in the Glaeser–Gottlieb model, individual preferences, and production and construction technologies are given by Cobb–Douglas functions. Individual

[18] The potential problem with this assumption is related to the fact that under the Soviet rule the spatial allocation of labor was guided by 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 a century of the post Soviet period a lot of people have moved to locations they preferred. See also Markevich and Mikhailova (2013)). For more details concerning the

(*footnote continued*)
migration patterns, see Appendix A.

Table 2
City age and wage in dynamics by age groups, 1991–2013.

City status	> 1982 (1)	Between 1953 and 1982 (2)	Between 1930 and 1953 (3)	Between 1917 and 1930 (4)	Between 1703 and 1917 (5)	Before 1703 (6)
year × log (*age*$_1$)	0.0085***	0.0101***	0.0046***	0.0044***	0.0028**	0.0066***
	[0.0019]	[0.0010]	[0.0011]	[0.0016]	[0.0012]	[0.0015]
year	0.0175	0.0778***	0.2202***	0.2171***	0.2388***	0.1624***
	[0.0166]	[0.0108]	[0.0102]	[0.0173]	[0.0106]	[0.0149]
log (*age*$_1$)	− 24.9417***	− 53.0307***	− 97.0294***	− 91.6968***	− 87.7120***	− 62.5142***
	[8.8041]	[4.8275]	[3.6101]	[6.2418]	[3.0014]	[3.7762]
Observations	1,117	4,394	5,832	1,668	5,183	3,368
Number of cities	70	221	285	82	253	164

Note: The estimates were obtained using the Arellano-Bond estimator. GMM standard errors are in brackets. The specifications include the first lag of the 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 3
Age and wage in other post-Soviet countries.

	Ukraine, 2013			Belarus, 2014			Kazakhstan, 2013			
	simple	+ log pop.	+ regions	simple	+ log pop.	+ regions	simple	+ log pop.	+ regions	
Log age	−0.1096***	−0.1048***	−0.0893***	−0.0360	−0.0523**	−0.0622***	−0.0219	−0.1424**	−0.1264**	
	[0.0146]	[0.0138]	[0.0200]	[0.0329]	[0.0240]	[0.0206]	[0.0512]	[0.0580]	[0.0502]	
Log pop.		0.0648***	0.0534***		0.0742***	0.0694***		0.1120***	0.0993***	
		[0.0076]	[0.0064]		[0.0098]	[0.0094]		[0.0287]	[0.0242]	
Obs.	458	458	458	112	112	112	87	87	87	
R^2 adj.	0.173	0.268	0.372	0.016	0.383	0.483	0.002	0.173	0.610	
	Azerbaidzhan, 2011			Tadzhikistan, 2013			Lithuania, 2010		Moldova, 2014	
	simple	+ log pop.	+ regions	simple	+ log pop.	+ regions	simple	+ log pop.	simple	+ log dens.
Log age	0.0138	−0.0367***	−0.0349***	−0.1242*	−0.1539**	−0.0909	−0.0929	−0.2352**		
	[0.0240]	[0.0114]	[0.0124]	[0.0645]	[0.0561]	[0.0988]	[0.0869]	[0.0707]		
Log pop.		0.1054***	0.0582***		0.1209	0.0479		0.2390**		
		[0.0339]	[0.0178]		[0.0704]	[0.2530]		[0.0915]		
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	17	17	17	10	10	32	26
R^2 adj.	0.016	0.331	0.592	0.220	0.202	0.063	0.089	0.471	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 on-line Appendix D.7. All the income variables are annual ones, except for Moldova for which the income variable is a monthly one for September.

firms and developers face a constant scale effect, but at the level of a location there is a diminishing scale effect due to fixed nontraded capital.

We extend the Glaeser–Gottlieb model to account for the role of natural resources. We treat the latter as those present in a spatial unit and extracted to be used as an input in the local production of the composite good.[19] This extension is motivated by the fact that it is extracted resources, rather than just those present in an area, that are the major force behind the relative performance of various spatial units of Russia. Besides, the local resource abundance can result either in the more available input for the production of the finished goods, or the additional income from the resource export. Whatever particular use of the local resource, the ultimate benefit from the resource stocks takes the form of higher consumption of the finished good. Therefore we add a separate extraction sector which supplies locally the input for the composite good production.

The functions contain four inputs: labor, traded capital, natural resources, and nontraded capital, among which all but the last are chosen at the local level. Appendices B.1–B.5 contain the simple 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 the regressions of the endogenous variables on city age. Solving the system A1, A4, A7, and A9 in Appendix B in logarithms for the exogenous variables and substituting the coefficients into the solution one obtains the relationships between city age and the exogenous variables:[20]

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

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

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

$$\lambda_B = b_R - b_N(\nu+\xi) - b_W\nu. \tag{4}$$

λ_T, λ_A, λ_θ, and λ_B are the city age coefficients from the linear functions of the production, construction, consumption, and extraction amenities. b_N, b_W, b_P, and b_R are the city age coefficients from the linear

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

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

regressions of labor force, wage, housing price, and natural resource use.[21] α denotes the housing elasticity of utility, β, γ and δ are labor, capital, and resource elasticities of the composite good production, ϵ and η, and ν and ξ are the respective parameters for labor and capital in the construction and extraction functions. For more details concerning the model and the parameters, see Appendices B.1–B.4.

Having specific values for 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 to determine the sources of the negative age-income correlations coupled with the relationships between city age and other endogenous variables.

As in Glaeser and Gottlieb (2009), the same system can be considered dynamically by first differencing the Eqs. A1, A4, A7, and A9. The respective solution of the first-difference equation system for the exogenous variables will be the same as those for the static model except that the coefficients from the static regressions are replaced by those from the first-difference 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, by adding 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 than 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[22] which provide them with different consumption amenities. Finally, the construction sector produces two kinds of housing with different productivity parameters.

This version of the model is derived in Appendix B.7. The solution of the system A15–A17, and A19 for the exogenous variables gives the following relationships between the coefficients of the regressions of the exogenous variables on a variable and the regression coefficients of the endogenous variables on the same variable

$$\lambda_{\Lambda} = (b_{\iota} + b_{\omega})(1 - \epsilon - \eta) - b_{\pi}, \tag{5}$$

$$\lambda_{\psi} = b_{\iota}(1 - \sigma) + b_{\omega}, \tag{6}$$

$$\lambda_{\vartheta} = b_{\pi}\alpha - b_{\omega}, \tag{7}$$

$$\lambda_{\phi} = b_{\rho}(1 - \varsigma(\nu + \xi)). \tag{8}$$

λ_{Λ}, $\lambda_{\psi'}$, λ_{ϑ}, λ_{ϕ} are the coefficients for the exogenous variables, namely for the relative construction amenities for high quality housing, the productivity parameter for skilled labor, the relative consumption amenities for skilled labor, and the relative extraction amenities of the exhaustible resource. b_{ι}, b_{ω}, b_{π}, and b_{ρ} are the coefficients for the endogenous variables, namely for the relative number of skilled workers, their relative wage, the relative price of high quality housing, and the relative output of the exhaustible resource. σ and ς are the substitution parameters for skilled and unskilled workers, and for exhaustible and renewable resources, respectively. For more details concerning this version of the model, see Appendix B.7.

This system allows one to readily move to the dynamic version by substituting the coefficients from the respective first-difference equations, as in the previous subsection.

[21] The resultant linear combination in 3 is the same as that in Glaeser (2008, p. 55). The additional coefficient defined in 4 follows from the inclusion of the new input in the production function and the addition of the extraction sector.

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

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 a more skilled workforce. This could be the result of a premium 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 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 should also be offset by higher wages.

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

4. Results and discussion

4.1. Econometric specifications

To test the hypotheses, a number of estimates are needed. To obtain the coefficients 1–4 and 5–8, based on the two versions of the model, we run a number of regressions. As dependent variables in the former case we use the logarithms of labor, wage, housing price, and resource 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 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 commercial 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 regional level.

To determine the relationships between city age and the change of all these variables, we run the dynamic specifications as in Glaeser and Gottlieb (2009) except that we use all the available cross-sections between the earliest and latest years. For obtaining the coefficients in the first-difference equations we run the panel regressions for 1996–2013. For the relative measures we estimated similar panel regressions for 2005–2009. The respective panel specifications are as follows:

$$y_{it} = b_1 year_t \times \log(age_i) + b_2 y_{it-1} + \mathbf{byr}_t + \alpha_i + \varepsilon_{it} \tag{9}$$

where y_{it} is the logarithm of the 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

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	−0.0806	−0.1741***	−0.0396**	−0.2879***
	[0.0609]	[0.0254]	[0.0161]	[0.0983]
Observations	930	930	927	921
R^2adj	0.377	0.342	0.122	0.433
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	−0.0188	−0.0398***	−0.0003	−1.0248*
	[0.0137]	[0.0100]	[0.0062]	[0.5447]
Observations	927	929	653	929
R^2adj	0.032	0.147	0.014	0.021

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

city's fixed effect, and ε_{it} is the *i*th city's idiosyncratic characteristics in the *t*th year. Assuming that the dependent variable lag y_{it-1} is a close covariate of many important urban characteristics it is included as a control variable. These regressions are estimated by the Arellano–Bond estimator. Like the respective specifications in Glaeser and Gottlieb (2009), the interaction term is the variable of interest.

4.2. Estimation results

Initial estimates come from sample statistics for the age groups of the cities (for the statistics and their discussion, see on-line Appendix D.4). Relating these statistics to the hypotheses to be tested, higher incomes in newer cities are likely to result from productivity advantages related to the higher level of human capital and/or resource abundance.

Panel A of Table 4 contains the regression estimates for the simple static model. The variable of interest is significant in columns 2–4. 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 0.29%, respectively. Age does not make any significant difference in terms of city size.[23]

Panel regression estimates to establish the comparative dynamics of these variables across different ages are presented in Panel A of Table 5. The results are significant for the variable of interest. Now wage, housing price, and resource variables go in the opposite direction compared to the static estimates. Like the results in Table 2 for age subgroups, for the whole population 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. City age predicts the opposite dynamics for income and population. These dynamics coupled with the preceding static results may be explained by in-migration to relatively rich new cities, which do not feature the most rapid income growth.[24] As follows from the spatial

Table 5
City age and change of the endogenous variables.

Panel A	Dependent variable, 1996–2013			
	Log labor (1)	Log wage (2)	Log housing price (3)	Log stocks (4)
year × log (*age*)	−0.0008**	0.0049***	0.0013***	0.0035**
	[0.0004]	[0.0003]	[0.0004]	[0.0018]
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 × log (*age*)	−0.0029***	0.0015***	0.0003	0.0830*
	[0.0005]	[0.0005]	[0.0011]	[0.0471]
Observations	7,506	3,215	4,756	3,225

Note: The estimates were obtained using the Arellano–Bond estimator. GMM standard errors 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

equilibrium model, these dynamics in older cities should go 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, feature higher relative prices for high quality housing, and use relatively more exhaustible resources compared to their use of renewable resources. The spatial equilibrium approach suggests that higher wages for skilled workers in newer cities may be offset by prices for high quality housing and 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, all the estimates except for column 3 are significant and with the same signs, which is also in line with the logic of the model. 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. These comparative dynamics may have reflected a change for the worse in older cities in their consumption amenities from the standpoint of skilled workers.

4.3. Theory-based interpretation

These results coupled with the spatial equilibrium model and the parameter estimates let one calculate the effects of city age on the exogenous variables, which may underlie the relationships of interest. Most parameter values are obtained estimating the production function on the same dataset for Russian cities (for data and estimation results, see on-line Appendix H). The remaining parameters are borrowed from the literature. The parameter values used in the calculations are in Table A3 in Appendix.

All the calculations are presented in Table 6. Column 1 contains the values of the coefficients λ_T, λ_A, λ_θ, and λ_B calculated based on the results of Table 4 in Panel A and formulas 1–4. As follows from these values, newer cities feature higher production, construction, and extraction amenities, and lower consumption amenities. This means that newer cities impose lower production costs on firms, due to better production and extraction conditions, and lower construction costs on developers, but provide their inhabitants with less pleasant living conditions.[25] Column 2 contains the coefficient values for the first-

[23] This result is at odds with the sample statistics for Russian cities and the existing evidence concerning the age-size correlations for developed countries (Giesen and Suedekum, 2014; Dobkins and Yannis, 2001; Michaels et al., 2012). However, as mentioned, the data for Russian cities are in line with these results in relatively young subsamples. For the whole population of cities the correlation holds 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.

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

[25] 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" .

Table 6
City age and the exogenous 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	−0.1266	0.004	−0.0398	0.0001
A/Λ	−0.1176	0.003	−0.006	−0.0002
θ/ϑ	0.17	−0.0048	0.0398	−0.0015
B/ϕ	−0.1588	0.0005	−0.5893	0.0477

Note: The coefficients in columns 1 and 2 for $\log(T)$, $\log(A)$, $\log(\theta)$, $\log(B)$ were calculated according to 1–4. The coefficients in columns 3 and 4 for $\log(\psi)$, $\log(\Lambda)$, $\log(\vartheta)$, $\log(\phi)$ were calculated according to 5–8.

difference version used to explain the income convergence observed for cities with different ages. All the values are 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 production, construction, and extraction amenities. This suggests that within any fixed group new cities are no longer such unpleasant places compared to older cities, while the efficiency differentials changed in favor of older cities. Smaller differences between the cities in all the amenities should have resulted in smaller real income differentials.

The results based on formulas 5–8 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 better relative conditions for extracting exhaustible resources in newer cities. The additional sources of the productivity advantages of new cities 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 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 4) suggests that the attractive and non-attractive characteristics of the new cities make skilled workers indifferent between newer and older cities.

Coefficients in column 4 capture the comparative dynamics of the relative measures. The values mostly have signs opposite to column 3. This also suggests convergences in 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 and construction amenities grew more rapidly in newer cities. Thus, from a skilled worker's perspective, newer cities became better for life and cheaper. 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 favor 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, a higher share of exhaustible resources, and a more efficient use of skilled workers. The productivity-related real wage premium should attract additional population, while the higher construction productivity should make their housing more available, which reinforces their population growth. The latter effect is weakened by the consumption disamenities because these should discourage people from living in newer cities. Therefore, the higher productivities in the production and extraction sectors do not fully transmit to higher housing prices.[26] 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 is − 0.59), their consumption disamenities (0.17), and the general resource abundance (− 0.16).

The convergent dynamics are especially fast for the share of exhaustible resources and the general resource stocks as seen in the change in the relative extraction amenities (0.048), in the general extraction amenities (0.001), and in consumption amenities (− 0.002). These can result from resource exhaustion or human activities to improve living conditions in newer cities.[27] Other urban characteristics, including general and relative consumption and construction amenities, and the return on human capital, show convergence too. Newer cities are better in their production (− 0.13) and construction amenities (− 0.12) and in their return on human capital (− 0.04), but their advantages became smaller with time.

The same results as those presented in Tables 4–6 are obtained for a subsample of resource extracting cities and for the remaining cities. For the calculations for the subsamples, see Table A4 in Appendix C. Briefly, more substantial relationships between city age and resource and consumption amenity variables are shown by resource extracting cities. For the other effects, an important difference between the groups is that resource extracting cities display convergence for all but one amenity variables, while the remaining cities show divergence for more variables.

4.4. Robustness check

To make sure the conclusions fit the data, we estimate a number of regressions using proxies for the exogenous variables. Like Glaeser (2008), we use the 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] A warmer climate, lower relative emissions, and higher theater availability indicate better consumption amenities. As a proxy for production, construction, and extraction amenities we use the residuals from the respective production functions.[29] The conclusions based on the simple static model predict that newer cities should be less pleasant places to live, but richer in resources. The robustness check confirmed that the model-based conclusions fit the data. Table 7 shows that newer cities are colder, dirtier, and feature fewer theater attendances, while their productivities in all the three sectors are higher. This should make their real wages higher for productivity- and amenity-related reasons.

4.5. 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

[26] 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).

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

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

[29] For details, see on-line Appendix H.

Table 7
City age and proxies for the exogenous variables, 2013.

	Temperature in January	Relative emissions	Theater attendances per capita	log (T)	log (A)	log (B)
	(1)	(2)	(3)	(4)	(5)	(6)
Log age	1.6498***	−32.4556***	0.0297***	−0.2798***	−0.6350***	−1.0827***
	[0.2594]	[12.3827]	[0.0061]	[0.1046]	[0.1716]	[0.1995]
Observations	1,075	1,075	1,075	1,028	1,028	1,028
R^2	0.0347	0.0106	0.0201	0.0075	0.0127	0.0320

Note: Robust standard errors obtained by the sandwich estimator are in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Schumpeterian creative destruction when applied to spatial input allocation and its rewards (Schumpeter, 1942; Hounshell, 1984; Florida, 1996). As 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 population and firms 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,[30] 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, their relative wages diminished and their populations declining.[31] Similar tendencies are seen in the oil and gas industry in Russia. Between 2005 and 2013 a number of oil and gas extracting cities experienced declining resource output per worker along with decreasing populations and relative wages.[32] These locations are mostly new cities paying a substantial wage premium, but their income advantages have diminished with time.

From a more general perspective, this 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 newly-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 a number of Russian cities (see, e.g., Kluchevsky, 1911; Lappo, 1998; Lappo, 2006).[33]

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. 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. 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 crowding effect, and, in fact, over the same period manufacturing production did increase (Dobrynskaya and Turkisch, 2009). While in Russia city age predicts relative resource scarcity and an 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 a 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. City size is a strong predictor of wage. Thus, as in the US, the older cities in Russia tend to be bigger and have an 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 weaker agglomeration forces. To determine the underlying relationships between city age and other

[30] For example, Usolye-Sibirskoye, Solvychegodsk, Solikamsk, Sol-Iletsk.

[31] The examples of such locations are Anzhero-Sunzhensk, Gukovo, Prokopievsk, Osinniki, Donetsk (a Russian town unlike the Ukrainian city with the same name), Zverevo, Novoshakhtinsk, and Kizel. The former three 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"), the latter is an example of dying town due to its closed collieries. Outside Russia an example of a location dying for the same reason is Ukrainsk in Ukraine.

[32] These include cities such as Pokachi, Uray, Pyt-Yakh, Nyagan.

[33] 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 the development of ones already controlled. Over the Soviet era the state used the vast space to retreat during war or to deal with economic difficulties in peace time.

urban characteristics, we use an extended version of the Glaeser–Gottlieb spatial equilibrium model. The model assumes the spatial no-arbitrage condition for population and firms. Based on these assumptions the equilibrium conditions in consumer and housing markets were derived with endogenously determined population, wage, housing price, and resource use. 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 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.

City age is linked with productivity- and amenity-related urban characteristics. New cities are more productive; these advantages of new cities result from their production amenities, higher return on skills, more abundant natural resources and the higher shares of exhaustible resources in their resource stocks. New cities are less pleasant places to live in. Higher real wages in new cities are a 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, but as the territory develops its disadvantages in consumption amenities also tend to diminish. From a broader perspective, this pattern corresponds to the way the Russian population and state have 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 to Vera Danilina, Sergey Izmalkov, Sergey Kokovin, Alexey Korolev, Grigory Kosenok, Tatiana Mikhailova, Emily Oster, Jens Suedekum, Jacques François Thisse, Stepan Jurajda, Dao-Zhi Zeng, Ives Zenou for valuable comments. 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. Migration patterns in the Soviet and post Soviet periods

To make sure that the pattern of internal migration observed in the post Soviet period differs from that typical for the Soviet period, we run a number of regressions of net migration per capita. As regressors we include the logarithms of latitude, longitude, wage, and price for the consumer basket, and the relative emissions. We treat them as measures of the modesty of climate and distance from densely populated regions, the level of life, and the quality of the natural environment, i.e., the characteristics which an individual takes into account when thinking about moving. We suppose that in the Soviet period people not having a choice where to live had to move to distant places, while in the post Soviet period people move to more attractive locations. Table A1 presents the estimation results.

The available data makes it possible to contrast the periods only for the effect of geographical coordinates. The results show that in the Soviet period people tended to move northward and eastward, and in the post Soviet period the directions of mass migration were reversed. The results also indicate that in the post Soviet years the directions of the internal migration are determined by wage, consumer prices, and the quality of the natural environment. In line with the previous evidence, the results suggest that nowadays people move motivated by rational considerations rather than orders from above.

Table A1
Net migration to population ratio.

	Soviet period		Post Soviet period					
	1980	1986	1992	1996	2000	2005	2010	2013
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log latitude	0.0310*	0.0846***	−0.0324***	−0.0201***	−0.0124***	−0.0161***	−0.0222***	−0.0311***
	[0.0188]	[0.0241]	[0.0093]	[0.0047]	[0.0045]	[0.0054]	[0.0046]	[0.0067]
Log longitude	0.0097**	0.0101***	−0.0087***	−0.0108***	−0.0060***	−0.0051***	−0.0073***	−0.0109***
	[0.0043]	[0.0033]	[0.0015]	[0.0011]	[0.0009]	[0.0007]	[0.0013]	[0.0011]
Relative emissions			0.0000***	−0.0000	−0.0000	−0.0000	−0.0000*	−0.0000*
			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Log wage			−0.0022	−0.0001	0.0028***	0.0042***	0.0111***	0.0173***
			[0.0014]	[0.0009]	[0.0009]	[0.0014]	[0.0030]	[0.0024]
Log consumer price						−0.0051*	−0.0134***	−0.0203***
						[0.0029]	[0.0035]	[0.0059]
Observations	387	321	948	1,023	1,035	1,052	1,049	1,042
R^2 adj.	0.0241	0.0975	0.114	0.133	0.0672	0.0710	0.0725	0.147

Note: Robust standard errors obtained by the sandwich estimator are in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix B. The spatial equilibrium model

B1. Consumer preferences and demand for housing

An individual should choose the optimal combination of housing H and composite good (other goods), which is used as a numeraire, given the locally specific wage W, and price of housing P. 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:

$$\overline{U} = k_1\theta WP^{-\alpha}. \tag{A1}$$

where $k_1 = \alpha^{\alpha}(1-\alpha)^{(1-\alpha)}$. 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:

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

This will be used when solving for equilibrium housing output.

B2. Production

The composite good is produced using Cobb–Douglas technology as follows

$$TN^{\beta}K^{\gamma}R^{\delta}Z^{1-\beta-\gamma-\delta} \tag{A3}$$

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

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. We assume that, unlike labor and capital, natural resources do not move, so that their price μ is an exogenous variable, and firms face different resource costs at various locations depending on the local resource output.[34] The profit equality across space suggests that the input prices and production amenities offset each other.

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

FOCs for labor, capital, and resources give the inverse demand functions for labor and natural resources:

$$W = (k_2 TZ^{1-\beta-\gamma-\delta}N^{-1+\gamma+\beta}R^{\delta})^{\frac{1}{1-\gamma}}, \tag{A4}$$

$$\mu = (k_3 TZ^{1-\beta-\gamma-\delta}N^{\beta}R^{-1+\gamma+\delta})^{\frac{1}{1-\gamma}}, \tag{A5}$$

where $k_2 = \beta^{1-\gamma}\gamma^{\gamma}$ and $k_3 = \delta^{1-\gamma}\gamma^{\gamma}$. Since, unlike the other inputs, capital is priced equally across space, the demand for capital is eliminated from the system. The labor demand function A4 is actually the equilibrium condition in the composite good market.[35]

B3. Construction sector and equilibrium housing price

The construction sector uses Cobb-Douglas technology too, but with different parameters and using only labor and capital as its inputs

$$AN^{\epsilon}K^{\eta}L^{1-\epsilon-\eta} \tag{A6}$$

where A is the productivity of the construction sector, and L is the nontraded capital of the specific form required by the construction sector,[36] which suggests a constant scale effect at the firm level and a decreasing scale effect at the location level. The constant $AL^{1-\epsilon-\eta}$ reflects the construction amenities of a place.

As the construction sector pays the same prices for inputs as the production sector, equal profit across space and sectors should make input prices correspond to construction amenities. A developer maximizes profit $PAN^{\epsilon}K^{\eta}L^{1-\epsilon-\eta} - WN - K$.

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

$$\left(k_4\frac{AL^{1-\epsilon-\eta}P^{\epsilon+\eta}}{W^{\epsilon}}\right)^{\frac{1}{1-\epsilon-\eta}}$$

where $k_4 = \epsilon^{\epsilon}\eta^{\eta}$. Using the demand function A2 one has the equilibrium housing price

[34] This assumption implies that the resource price is inversely related to the remaining stock of the respective resource, which is in line with the exhaustible resource literature originating from Hotelling (1931).

[35] This can be readily checked if one derives the supply function using FOCs for capital, labor, and resources, equates this to the demand $(1-\alpha)WN$, and solves the resulted equilibrium condition for wage.

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

$$P = k_5 \frac{N^{1-\epsilon-\eta} W^{1-\eta}}{AL^{1-\epsilon-\eta}} \tag{A7}$$

where $k_5 = \alpha^{1-\epsilon-\eta}\epsilon^{-\epsilon}\eta^{-\eta}$.

B4. Extraction sector and equilibrium resource output

The extraction sector uses Cobb-Douglas technology as follows

$$BN^{\nu}K^{\xi}M^{1-\nu-\xi} \tag{A8}$$

where B is the productivity shifter in the extraction sector, and M is the nontraded capital required by the extraction sector. The constant $BM^{1-\nu-\xi}$ reflects the location-specific extraction amenities. An extractive firm maximizes profit $\mu BN^{\nu}K^{\xi}M^{1-\nu-\xi} - WN - K$. Substituting FOCs for the inputs in the extraction function and solving for the resource price gives the inverse resource supply function:

$$\mu = \left(k_6 \frac{R^{1-\nu-\xi} W^{\xi}}{BM^{1-\nu-\xi}}\right)^{\frac{1}{\nu+\xi}}$$

where $k_6 = \nu^{-\nu}\xi^{-\xi}$. Using the inverse demand function for the resource A5 one has the equilibrium resource output:

$$R = \left(k_7 \frac{(TZ^{1-\beta-\gamma-\delta})^{\nu+\xi}(BM^{1-\nu-\xi})^{1-\gamma} N^{\beta(\nu+\xi)}}{W^{\xi(1-\gamma)}}\right)^{\frac{1}{1-\gamma-\delta(\nu+\xi)}} \tag{A9}$$

where $k_7 = \gamma^{\gamma(\nu+\xi)}\delta^{(1-\gamma)(\nu+\xi)}\nu^{\nu(1-\gamma)}\xi^{\xi(1-\gamma)}$.

B5. The effect of city age on the exogenous variables

The Eqs. A1, A4, A7, and A9 contain spatial equilibrium conditions for labor, housing, and natural resources, where *N, W, P*, and *R* are the endogenous variables and $TZ^{1-\beta-\gamma-\delta}$, $AL^{1-\epsilon-\eta}$, θ, and $BM^{1-\nu-\xi}$ are the exogenous variables. Thus, we have four equations, four endogenous variables, and four exogenous variables. Taking the logarithms of the system A1, A4, A7, and A9 and solving it for the endogenous variables we obtain linear equations with the coefficient matrix as follows:

$$\mathbf{c} = \frac{1}{k_8}\begin{pmatrix} 1-\alpha(1-\eta) & \alpha(1-\gamma-\delta\nu) & 1-\gamma-\delta\nu & \delta(1-\alpha(1-\eta)) \\ \alpha(1-\epsilon-\eta) & \alpha(\beta+\gamma+\delta(\nu+\xi)-1) & \beta+\gamma+\delta(\nu+\xi)-1 & \alpha\delta(1-\epsilon-\eta) \\ 1-\epsilon-\eta & \beta+\gamma+\delta(\nu+\xi)-1 & (\beta+\delta\xi)(1-\eta)-\epsilon(1-\gamma-\delta\nu) & \delta(1-\epsilon-\eta) \\ \nu(1-\alpha\epsilon)+\xi(1-\alpha(1-\eta)) & \alpha(\beta\nu+\xi(1-\gamma)) & \beta\nu+\xi(1-\gamma) & (1-\gamma)(1-\alpha\epsilon)-\beta(1-\alpha(1-\eta)) \end{pmatrix} \tag{A10}$$

where $k_8 = (1-\alpha\epsilon)(1-\gamma-\delta\nu) - (\beta+\delta\xi)(1-\alpha(1-\eta))$. The signs of the coefficients in A10 are in line with the theory.[37] The production amenities (the first column of **c**) positively affect all the endogenous variables. The construction amenities (the second column of **c**) lower construction costs, which makes housing prices lower (the negative sign of c_{32}). This attracts more population (the positive sign of c_{12}), which decreases wages (c_{22}) and increases the use of natural resources (c_{42}). The consumption amenities (the third column of **c**) attract more population (c_{13}), which is offset by lower wages (c_{23}) and makes housing prices higher (c_{33}). More people suggest more use of resources (c_{43}). The resource extraction amenities (the fourth column of **c**) make resources available (c_{44}), which induces the labor demand (c_{14}). The latter in turn increases wages (c_{24}) and housing prices (c_{34}).

B6. The dynamic system

The first-differences of Eqs. A1, A4, A7, and A9 are as follows

$$\begin{aligned}
\frac{P_{t+1}}{P_t} &= \left[(1+g_\theta)\frac{W_{t+1}}{W_t}\right]^{\frac{1}{\alpha}}, \\
\frac{W_{t+1}}{W_t} &= \left[\frac{(1+g_T)\left(\frac{R_{t+1}}{R_t}\right)^{\delta}}{\left(\frac{N_{t+1}}{N_t}\right)^{1-\beta-\gamma}}\right]^{\frac{1}{1-\gamma}}, \\
\frac{P_{t+1}}{P_t} &= \frac{\left(\frac{N_{t+1}}{N_t}\right)^{1-\epsilon-\eta}\left(\frac{W_{t+1}}{W_t}\right)^{1-\eta}}{1+g_A}, \\
\frac{R_{t+1}}{R_t} &= \left[\left((1+g_T)\left(\frac{N_{t+1}}{N_t}\right)^{\beta}\right)^{\nu+\xi}\left(\frac{(1+g_B)}{\left(\frac{W_{t+1}}{W_t}\right)^{\xi}}\right)^{1-\gamma}\right]^{\frac{1}{1-\gamma-\delta(\nu+\xi)}}.
\end{aligned} \tag{A11}$$

where g_θ, g_T, g_A, g_B are the exogenous growth rates of the location-specific amenities in consumption, production, construction, and resource

[37] The items of the respective block of the matrix correspond to those in Table 2 in Glaeser and Gottlieb (2009, p. 994).

extraction, respectively. From the standpoint of an individual, a change of income can be offset by a change of current housing costs, rather than housing price. However, assuming the constant expected growth of housing prices $\frac{P_{t+1}}{P_t}$ this can serve as a proxy for the growth rate of rental costs.[38]

B7. Two kinds of human capital and resources

The consumer problem is the same as in A1, but now individuals are divided into skilled and unskilled workers. These differ from each other by their wage, type of housing, and respective consumption amenities enjoyed. The indirect utility function for both types of human capital is defined as in A1. Hence, the relative indifference condition of the skilled workers is as follows:

$$\frac{\overline{U}_S}{\overline{U}_U} = \frac{\theta_H W_S}{\theta_L W_U}\left(\frac{P_H}{P_L}\right)^{-\alpha} \tag{A12}$$

where indices S and U refer to skilled and unskilled workers, and indices H and L refer to the high and low quality of housing and consumption amenities. As in A2, we get the demand functions for two kinds of housing, from which we have the relative demand for high level housing:

$$\frac{D_H}{D_L} = \frac{W_S N_S}{W_U N_U}\left(\frac{P_H}{P_L}\right)^{-1}. \tag{A13}$$

The construction sector builds the two kinds of housing using the technologies as follows

$$\begin{aligned} H_H &= A_H(\psi N_S^\sigma + N_U^\sigma)^{\frac{\epsilon}{\sigma}} K^\eta L_H^{1-\epsilon-\eta} \\ H_L &= A_L(\psi N_S^\sigma + N_U^\sigma)^{\frac{\epsilon}{\sigma}} K^\eta L_L^{1-\epsilon-\eta} \end{aligned} \tag{A14}$$

where ψ is the productivity parameter, N_S and N_U are skilled and unskilled labor. To derive the equilibrium price of high (low) quality housing, we use the construction functions A14, FOCs for capital and the two kinds of labor. To simplify the system, treat $\psi N_S^\sigma + N_U^\sigma$ as an endogenous variable and proceed to the three equation system in which the remaining two endogenous variables are H_H (H_L) and K. Solving the system we have the high (low) quality housing supply function. The relative high quality housing supply function is as follows

$$\frac{H_H}{H_L} = \left(\frac{A_H L_H^{1-\epsilon-\eta} P_H^{\epsilon+\eta}}{A_L L_L^{1-\epsilon-\eta} P_L^{\epsilon+\eta}}\right)^{\frac{1}{1-\epsilon-\eta}}.$$

Finally, using the relative demand A13, we have the equilibrium relative price of high quality housing:

$$\pi = \frac{(\iota\omega)^{1-\epsilon-\eta}}{\Lambda}. \tag{A15}$$

where $\pi \equiv \frac{P_H}{P_L}$, $\iota \equiv \frac{N_S}{N_U}$, $\omega \equiv \frac{W_S}{W_U}$, and $\Lambda \equiv \frac{A_H L_H^{1-\epsilon-\eta}}{A_L L_L^{1-\epsilon-\eta}}$.

Then, we substitute the equilibrium relative price function A15 into the relative indifference condition A12 and take the respective ratio between wages of the two kinds of workers to obtain

$$\omega = \left[\frac{\iota^{\alpha(1-\epsilon-\eta)}}{\vartheta\Lambda^\alpha}\right]^{\frac{1}{1-\alpha(1-\epsilon-\eta)}} \tag{A16}$$

where $\vartheta \equiv \frac{U_U\theta_H}{U_S\theta_L}$.

The technology in the production sector is $T(\psi N_S^\sigma + N_U^\sigma)^{\frac{\beta}{\sigma}} K^\gamma (R_e^\varsigma + R_r^\varsigma)^{\frac{\delta}{\varsigma}} Z^{1-\beta-\gamma-\delta}$ where R_e and R_r denote exhaustible and renewable resources, respectively. FOCs for two kinds of labor and resources give the relative inverse demand functions for labor and resources

$$\omega = \psi\iota^{\sigma-1}, \tag{A17}$$

$$\frac{\mu_e}{\mu_r} = \rho^{\varsigma-1} \tag{A18}$$

where $\rho \equiv \frac{R_e}{R_r}$. The extractive sector uses technologies $B_e(\psi N_S^\sigma + N_U^\sigma)^{\frac{\nu}{\sigma}} K^\xi M_e^{1-\nu-\xi}$ and $B_r(\psi N_S^\sigma + N_U^\sigma)^{\frac{\nu}{\sigma}} K^\xi M_r^{1-\nu-\xi}$ for exhaustible and renewable resources, respectively. Then proceeding in a similar way as we do in the case of housing supply we have the inverse relative resource supply function:

$$\frac{\mu_e}{\mu_r} = \left(\frac{\rho^{1-\nu-\xi}}{\phi}\right)^{\frac{1}{\nu+\xi}}$$

where $\phi \equiv \frac{B_e M_e^{1-\nu-\xi}}{B_r M_r^{1-\nu-\xi}}$. Then using the relative inverse demand function A18 we have the equilibrium relative output of exhaustible resource:

$$\rho = \phi^{\frac{1}{1-\varsigma(\nu+\xi)}}. \tag{A19}$$

B8. Do the model predictions fit the data?

The model suggests (e.g., the coefficient matrix A10) 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 run regressions of the endogenous variables on the proxies for

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

Table A2
Exogenous and endogenous variables, 2013.

	Log population	Log real wage	Log stocks
	(1)	(2)	(3)
Temperature in January	0.0191***	− 0.0120***	0.0096
	[0.0037]	[0.0010]	[0.0092]
Relative emissions	− 0.0001	0.0001***	0.0002
	[0.0001]	[0.0000]	[0.0002]
Theater attendances per capita	2.2774***	− 0.0466	2.9119***
	[0.3399]	[0.0488]	[0.4375]
Log extractive output per worker	0.0311***	0.0082***	0.0669***
	[0.0051]	[0.0013]	[0.0095]
Log manufacturing output	0.1428***	0.0063**	0.3750***
	[0.0097]	[0.0028]	[0.0268]
Log real wage	0.7762***		
	[0.1376]		
Log population		0.0647***	
		[0.0103]	
Observations	1,046	1,046	1,036
R^2 adj.	0.608	0.330	0.511

Note: Robust standard errors obtained by the sandwich estimator are in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

consumption amenities and resource use. The regressions of log population, log real wage, and log stocks are estimated controlling for the log real wage, log population, and the two production variables, respectively. The results are in Table A2.

Higher temperatures in January and more theater availability are consistent with bigger cities, while higher temperatures 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 relative emissions and the theater attendance correlate with the resource stocks in line with the model, while temperature is insignificant.

B9. The parameter values

The parameter values we use in the calculations are in Table A3.

The only parameter for which we have not found any ready estimate, because of a lack of proper empirical evidence (Growiec and Schumacher, 2006), is ς. For this parameter we assume substitutability between 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.

Table A3
The parameters.

Parameter	Value	Source
α	0.103	The EU in the world (2015, p. 35)
β	0.7104	Table A5 in on-line Appendix F.2; Kuboniwa (2011, p. 8), Rõõm (2001, p. 10)
γ	0.0627	Table A5 in on-line Appendix F.2
δ	0.1269	Table A5 in on-line Appendix F.2
ϵ	0.7532	Table A5 in on-line Appendix F.2; Serebryakov (2000, p. 157)
η	0.0968	Table A5 in on-line Appendix F.2
ν	0.7413	Table A5 in on-line Appendix F.2
ξ	0.1087	Table A5 in on-line Appendix F.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)

Appendix C. City age and the exogenous variables by groups

The static results in columns 1 and 3 mostly have the same sign for both subsamples. The age effects on consumption amenities and the specific consumption amenities for 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 convergences in production and construction amenities for both subsamples, but among the resource cities these are faster.

All the effects for resource extracting cities shown in Panel A of Table 4 A have the same signs as those in Table 6 and have higher absolute values for the resource, consumption amenity, and human capital variables. Panel B shows that the remaining cities display either the same effects by sign but to a much lesser extent or the opposite effects as for the divergence in resource abundance and the return on skills.

Table A4
City age and the exogenous variables for resource and non-resource cities.

λ-s	(1)	(2)	(3)	(4)
Panel A: resource cities				
T/ψ	− 0.1045	0.0045	− 0.0478	0.0007
A/Λ	− 0.0773	0.004	− 0.0072	− 0.0001
θ/ϑ	0.197	− 0.0059	0.0478	− 0.0019
B/ϕ	− 0.1682	0.0041	− 1.811	0.0885
Panel B: non-resource cities				
T/ψ	− 0.1269	0.0015	− 0.0484	− 0.0017
A/Λ	− 0.1204	0.0005	− 0.0101	− 0.0005
θ/ϑ	0.14	− 0.0018	0.0298	0
B/ϕ	− 0.0924	− 0.0005	0	0

Note: See Table 6.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.jue.2017.12.003

References

Alexeev, M., Conrad, R., 2009. The elusive curse of oil. Rev. Econ. Stat. 91 (3), 586–598.
Altonji, J.G., Elder Todd, E., Taber Christopher, R., 2005. Selection on observed and unobserved variables: assessing the effectiveness of catholic schools. J. Pol. Econ. 113 (1), 151–184.
Andrienko, Y., Guriev, S.M., 2004. Determinants of interregional mobility in russia. Econ. Transit. 12 (1), 1–27.
Autor, D.H., Dorn, D., 2013. The growth of low-skill service jobs and the polarization of the US labor market. Am. Econ. Rev. 103 (5), 1553–1597.
Bariev, I.K., 2007. Koordinaty naselennykh punktov, chasovye poyasa i ismenenia ischisleniya vremeni. Spravochnik. Konek, Moscow.
Beaudry, P., Green, D.A., Sand, B.M., 2014. Spatial equilibrium with unemployment and wage bargaining. theory and estimation. J. Urban Econ. 79 (1), 2–19.
Behar, A., 2010. The elasticity of substitution between skilled and unskilled labor in developing countries is about 2. Int. Monetary Fund. Selected works.
Brown, D., Fay, M., Felkner, J., Lall, S.V., 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.
Carluer, F., 2005. Dynamics of russian regional clubs: the time of divergence. Reg. Stud. 39 (6), 713–726.
Combes, P.-P., Mayer, T., Thisse, J.-F., 2008. Economic Geography. The integration of regions and nations. Princeton University Press, Princeton.
Davis, D.R., David, E.W., 2002. Bones, bombs, and break points: the geography of economic activity. Am. Econ. Rev. 92 (5), 1269–1289.
Dobkins, L.H., Yannis, M.I., 2001. Spatial interactions among U.S. cities: 1900–1990. Reg. Sci. Urban Econ. 31 (6), 701–731.
Dobrynskaya, V., Turkisch, E., 2009. Is Russia sick with the Dutch disease? CEPII no. 20.
Dolinskaya, I., 2002. Transition and regional inequality in Russia: reorganization or procrastination?. IMF working Paper 02/169.
The EU in the world - 2015 edition, 2015. Luxembourg: Eurostat. Publications office of the European union.
Florida, R., 1996. Regional creative destruction: production organization, globalization, and the economic transformation of the midwest. Econ. Geogr. 72 (3), 314–334.
Ekonomika gorodov Rossii, 2016. GMZ rosstata. http://multistat.ru/?menu_id=9310004 (accessed February, 2016).
Fu, Y., Gabriel, S.A., 2012. Labor migration, human capital agglomeration and regional development in china. Reg. Sci. Urban Econ. 42, 473–484.
Giesen, K., Suedekum, J., 2014. City age and city size. Eur. Econ. Rev. 71, 193–208.
Glaeser, E., Kolko, J., Saiz, A., 2001. Consumer city. J. Econ. Geogr. 1, 27–50.
Glaeser, E.L., 2008. Cities, Agglomeration, and Spatial Equilibrium. Oxford University Press, Oxford and New York.
Glaeser, E.L., 2011. Triumph of the City. How our greatest invention makes us richer, smarter, greener, healthier, and happier. Penguin Press, New York.
Glaeser, E.L., Gottlieb, J.D., 2009. The wealth of cities: agglomeration economies and spatial equilibrium in the united states. J. Econ. Lit. 47 (4), 983–1028.
Glaeser, E.L., Kahn, M.E., 2001. Decentralized Employment and the Transformation of the American City. NBER Working Paper No. 8117.
Gollin, D., Jedwab, R., Vollrath, D., 2016. Urbanization with and without industrialization. J. Econ. Growth 21 (1), 35–70.
Growiec, J., Schumacher, I., 2006. On technical change in the elasticities of resource inputs // CORE discussion paper 2006/63.
Guriev, S., Vakulenko, E., 2012. Convergence between Russian Regions. CEFIR/NES Working Paper No 180.
Heleniak, T., 1999. Out-migration and depopulation of the russian north during the 1990s. Post Sov. Geogr. Econ. 40 (3), 155–205.
Herzfeld, T., 2008. Inter-regional output distribution: a comparison of russian and chinese experience. Post-Commun. Econ. 20 (4), 431–447.
Hotelling, H., 1931. The economics of exhaustible resources. J. Pol. Econ. 39 (2), 137–175.
Hounshell, D., 1984. From American System to Mass Production. Johns Hopkins University Press, Baltimore, pp. 1800–1932.
James, A., Aadland, D., 2011. The curse of natural resources: an empirical investigation of U.S. countries. Resour. Energy Econ. 33 (2), 440–453.
Kholodilin, K. A., Oshchepkov, A., Siliverstovs, B., 2009. The Russian regional convergence process: where does it go? KOF working paper no. 216.
Kluchevsky, V.O., 1911. History of Russia. Vol. I. J. M. Dent & Sons, Ltd, London.
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.
Lappo, G.M. (Ed.), 1998. Goroda Rossii. Enziklopediya. Bolshaya Rossiyskaya Enziklopediya, Moscow.
Lappo, G.M., 2006. Vozrast gorodov. Geographia 7–8.
Lee, S., Li, Q., 2013. Uneven landscapes and city size distributions. J. Urban Econ. 78 (1), 19–29.
Markevich, A., Mikhailova, T.N., 2013. Economic Geography of Russia. In: Alexeev, M., Weber, S. (Eds.), The Oxford Handbook of the Russian Economy. Oxford University Press, Oxford.
Michaels, G., Rauch, F., Redding, S.J., 2012. Urbanization and structural transformation. Q. J. Econ. 127 (2), 535–586.
Mikhailova, T., 2011. Looking for Multiple Equilibria in Russian Urban System. Working Paper No10/08E.
Moretti, E., 2011. Local labor markets. Handbook of Labor Economics. 4b. pp. 1237–1313.
Moretti, E., 2013. Real wage inequality. Am. Econ. J 5 (1), 65–103.
Oster, E., 2017. Unobservable selection and coefficient stability: theory and validation. J. Bus. Econ. Stat. Forthcoming
Rattsø, J.r., Stokke, H.E., 2014. Population divergence and income convergence: regional distribution dynamics for norway. Reg. Stud. 48 (11), 1884–1895.
Roback, J., 1982. Wages, rents, and the quality of life. J. Pol. Econ. 90 (6), 1257–1278.
Rodgers, A., 1974. The locational dynamics of soviet industry. Ann. Assoc. Am. Geogr. 64 (2), 226–240.
Schumpeter, J.A., 1942. Capitalism, Socialism, and Democracy. In: Harper & Row, New York.
Serebryakov, G., 2000. Russian Economy: Structural Change and Econometric Perspective. In: Welfens, Paul, J.J., Gavrilenkov, E. (Eds.), Restructuring, stabilizing and modernizing the new Russia. Economic and institutional issues. Springer, New York.
SSSR, 1987. Administrativno-territorialnoye delenie soyuznykh respublik. Izvestiya Sovetov Narodnych Deputatov, Moscow.