Journal of Urban Economics 96 (2016) 112–120

Contents lists available at ScienceDirect

Journal of Urban Economics

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

Wage growth, ability sorting, and location choice at labor-force entry: New evidence from U.S. Census data

Zhi Wang

School of Economics, Fudan University, No. 600 Guoquan Road, Shanghai, 200433, China

ARTICLE INFO

Article history:
Received 6 May 2013
Revised 30 June 2016
Available online 3 August 2016

JEL Classification:
J3
R1
R23

Keywords:
Labor-force entry
Wage growth
Selection
Urban wage premium

ABSTRACT

Using a sample of college graduates from the 2000U.S. Census, this paper estimates the relationship between the location of labor-force entry and wage growth. The results show faster wage growth among workers who spend their early working years in larger, more densely populated cities with more highly educated populations. The estimates suggest that, for an average-ability worker entering the labor force in a large city, the initial wage premium upon labor-force entry is 15.7 percent, and the wage-growth premium after 5 years is 7.2 percent. Individual ability boosts wage growth in large cities but not in small cities or rural areas. As a result, higher-ability college graduates sort themselves into larger cities when they enter the labor force, leading to an additional urban wage-growth gap of 3.2 percent after 5 years.

1. Introduction

Between 1995 and 2000, 66 percent of U.S. college graduates changed their metropolitan statistical areas (MSAs) of residence when they entered the labor force. These young, highly educated workers are especially likely to move to larger cities (Glaeser, 1999; Chen and Rosenthal, 2008), which have denser input-output linkages between buyers and suppliers, more efficient labor markets, and faster human capital accumulation (Duranton and Puga, 2004). The latter two advantages associated with city size may cause faster wage growth in large cities for young workers. Several recent empirical studies have investigated the effect of city population on wage growth using panel data.[1] For instance, using samples drawn from the National Longitudinal Survey of Youth 1979 (NLSY79) and the Panel Study of Income Dynamics (PSID), Glaeser and Mare (2001) suggest that a significant fraction of the urban wage premium accrues to workers over time and stays with them when they leave the city. Using the NLSY79, Baum-Snow and Pavan (2012) show that wage-growth effects explain 57 percent of the wage premium of the largest MSAs over small MSAs and rural areas for college graduates and 78 percent of this premium for high school graduates.

Limitations of the panel data sets like the NLSY79 and the PSID include that their sample size is relatively small and their geographic coverage is limited. In this paper, I develop a novel empirical strategy that allows me to estimate the relationship between the location of labor-force entry and wage growth using a sample of college graduates from the 2000 Census data. Taking advantage of the large sample of Census data, I estimate wage-growth differentials for 328 localities covering the full contiguous U.S. states. The first step of my analysis is to measure the initial wage levels and the rates of wage growth of college graduates in different U.S. cities. I use the cohort of the age group 22–26 to estimate the initial wage levels. Because the Census data do not include historical wages, I estimate the wage-growth rates using an older cohort of the age group 27–31 under the assumption that the initial wage levels of the two cohorts are the same.

Previous research suggests that high-ability workers may experience faster wage growth in larger cities (Baum-Snow and Pavan, 2012). This implies that the wage premium in large cities in my sample may be caused in part by ability sorting. The Census data, however, do not contain any measure of ability beyond education attainment. To overcome this problem, I create a migration-based proxy for ability. According to the migration literature, migrants to cities far from their home towns bear a higher permanent transportation cost to maintain social connections in their home towns. Therefore, any state that is far from a large city should have a smaller fraction of native-born workers who enter the labor force in large cities. When ability is normally distributed, the average ability of workers in any city who were born in same state can

E-mail address: wangzhi@fudan.edu.cn

[1] See, for example, Glaeser and Mare (2001), Wheeler (2006), Yankow (2006), Gould (2007), and Baum-Snow and Pavan (2012).

http://dx.doi.org/10.1016/j.jue.2016.07.002
0094-1190/© 2016 Published by Elsevier Inc.

be explicitly expressed as a decreasing function of the fraction of workers from that state who entered the labor force in a large city. Using workers' birth states and locations of labor-force entry, I construct ability proxies and estimate the effects of ability on wage growth and the extent of ability sorting. I also provide corroborating evidence from the NLSY79, which contains individual Armed Force Qualification Test (AFQT) scores that measure individual ability before labor-force entry.

The results indicate that wage growth is faster if a worker enters the labor force in a larger, more densely populated city with a more highly educated population. For an average-ability worker entering the labor force in a large city, the initial wage premium upon labor-force entry is 15.7 percent, and the wage-growth premium after 5 years is 7.2 percent. A one-standard-deviation increase in ability raises wage growth by 4.3 percent in large cities, whereas no evidence of such an ability effect is found in small cities or rural areas. As a result, higher-ability college graduates sort themselves into larger cities when they enter the labor force, leading to an additional urban wage-growth gap of 3.2 percent after 5 years. Overall, for the workers entering the labor force in large cities, the wage premium is about 26.1 percent after 5 years, and the sorting of ability accounts for roughly 10 percent of this wage gap.

My estimate of the overall wage premium is consistent with the existing literature (e.g., Glaeser and Mare, 2001; Baum-Snow and Pavan, 2012). Previous studies based on the NLSY79 have provided mixed evidence on the importance of ability sorting in the U.S. Yankow (2006) suggests that two thirds of the wage premium in urban areas with populations of more than 1 million can be explained by the sorting of higher unmeasured skills and ability into these localities. Glaeser and Mare (2001) show that controlling for individual AFQT does not significantly affect the magnitude of the urban wage premium. Baum-Snow and Pavan (2012) even find that sorting contributes negatively to the observed urban wage premium. Compared with the NLSY79, the Census data set is larger and more representative.

The remainder of this paper is organized as follows. Section 2 presents the model. Section 3 discusses the data and the empirical method. Section 4 presents the main results and specification checks. Section 5 provides corroborating evidence from the NLSY79 data. Section 6 concludes.

2. Model

There is a continuum of individuals, indexed by i, who are identical except for a_i, their inherent ability, and b_i, the state in which each was born. There are B states, and the ability of the individuals born in each state follows a standard normal distribution.

There are two periods, and every individual works in both periods. I call an individual in period 1 a young worker (Y), and an individual in period 2 an old worker (O). At the beginning of period 1, each individual chooses to enter the labor force in one of J locations, which, for convenience, I refer to as a "city."

An individual i who works in city k in period 1 receives a wage

$$w^Y_{i,k} = w_k, \tag{1}$$

where w_k is the initial wage level in city k. In addition to w_k, an old worker in city k who worked in city j in period 1 receives a wage increase $\Delta w_{i,j,k}$. His total wage is therefore equal to

$$w^O_{i,j,k} = w_k + \Delta w_{i,j,k}. \tag{2}$$

I assume that

$$\Delta w_{i,j,k} = g_j + \pi_j a_i + \varepsilon_{i,k}, \tag{3}$$

where g_j is the wage growth specific to period-one city j, π_j is the effect of ability on wage growth, and $\varepsilon_{i,k}$ is an individual-level random shock that is realized at the end of period 1. I assume that $\varepsilon_{i,k}$ is independently and identically distributed over i and over k.

For the sake of estimation, I classify J cities into two categories based on their population sizes; namely, the small-city category and the large-city category. I also assume that π_j depends only on whether j is large or small. Let π_L and π_S denote the effects of ability on wage growth in a large city and in a small city, respectively.[2] I assume that $\pi_L > \pi_S$ in order to capture the idea that the wage growth differential between higher-ability and lower-ability individuals is greater in large cities.

At the beginning of period 1, each individual i chooses a city j in which to work in period 1 in order to maximize his expected lifetime utility

$$V_{i,j} = w^Y_{i,j} + \eta_{b_i,j} + E\left[\max_{k\in\{1,\dots,J\}}\{w^O_{i,j,k} + \eta_{b_i,k}\}\right]. \tag{4}$$

In (4), $\eta_{b_i,j}$ and $\eta_{b_i,k}$ denote the part of worker i's utility that is affected by non-productive attributes of the city (e.g., consumption amenities, living costs). The term $\eta_{b_i,j}$ (similarly $\eta_{b_i,k}$) varies by b_i, the birth state of individual i. People living in cities far from their birth states may have to bear a permanent transportation cost to maintain social connections. In (4), the expectation is taken with respect to $\varepsilon_{i,k}$. Some individuals may move after observing $\varepsilon_{i,k}$ at the end of period 1.

Because $\varepsilon_{i,k}$ is independently and identically distributed over i and over k, the expected lifetime utility $V_{i,j}$ depends only on individual i's ability, birth state, and period-one city. I can therefore rewrite (4) as

$$V_{i,j} = w_j + g_j + \pi_j a_i + \eta_{b_i,j} + \psi_{b_i}, \tag{5}$$

where $\psi_{b_i} \equiv E[\max_{k\in\{1,\dots,J\}}\{w_k + \eta_{b_i,k} + \varepsilon_{i,k}\}]$. While I assume that $\eta_{b,j}$ is exogenous, in a spatial-equilibrium model (e.g., Roback, 1982), a large city's greater productivity (i.e., its higher w_j or g_j) will lead to a higher living cost (i.e., lower $\eta_{b,j}$), which will prevent low-ability workers from moving into that large city. Let $\rho(j) \in \{S, L\}$ denote the size of city j. For birth state b, there exists a unique selection threshold

$$a^*_b = -\frac{\max_{\{j:\rho(j)=L\}}\{w_j + g_j + \eta_{b,j}\} - \max_{\{j:\rho(j)=S\}}\{w_j + g_j + \eta_{b,j}\}}{\pi_L - \pi_S} \tag{6}$$

such that individuals with ability above a^*_b move to the "best city" in the large-city category (i.e., the city with the highest value of $\{w_j + g_j + \eta_{b,j}\}$ among large cities), while those with ability below a^*_b move to the "best city" in the small-city category. When there are multiple best cities in the same category, the individual will randomly select one.

Based on (3), the average wage growth conditional on birth state (b) and period-one city (j) is

$$E(\Delta w_{i,j,k}) = g_j + \pi_S E(a|a < a^*_{b_i}) I_{\rho(j)=S} + \pi_L E(a|a > a^*_{b_i}) I_{\rho(j)=L}, \tag{7}$$

where $I_{\rho(j)=S}$ is an indicator function that equals 1 if city j is small, and $I_{\rho(j)=L}$ is an indicator function that equals 1 if city j is large. Note that both $E(a|a < a^*_b)$ and $E(a|a > a^*_b)$ vary with birth state. This birth-state variation is key to the estimation of g_j, π_S, and π_L. Although ability is not directly observed, P_b, the fraction of individuals born in state b who live in a large city in period 1, is known.

[2] The estimation method requires that there be a sufficiently large number of individuals drawn from each birth state within each city category of identical ability effect on wage growth. Therefore, the most straightforward practice is to conduct an estimation for the case where there are two size categories. Technically, I can easily adjust the model to accommodate cases where there are multiple city categories, and these estimation results are available upon request.

Let Φ, Φ^{-1}, and ϕ denote the CDF, inverse of CDF and PDF of the standard normal distribution, respectively. Since we assume that ability follows a standard normal distribution

$$P_b = 1 - \Phi(a_b^*). \tag{8}$$

I can therefore construct the following proxy measures for $E(a|a < a_b^*)$ and $E(a|a > a_b^*)$:

$$\lambda_{b,S} \equiv E(a|a < a_b^*) = -\frac{\phi(\Phi^{-1}(1-P_b))}{1-P_b}, \tag{9}$$

$$\lambda_{b,L} \equiv E(a|a > a_b^*) = \frac{\phi(\Phi^{-1}(1-P_b))}{P_b}. \tag{10}$$

Substituting (7), (9), and (10) into (3), I obtain the wage-growth regression model

$$\Delta w_{i,j,k} = g_j + \pi_S \lambda_{b_i,S} I_{\rho(j)=S} + \pi_L \lambda_{b_i,L} I_{\rho(j)=L} + u_{i,j,k}, \tag{11}$$

where

$$u_{i,j,k} \equiv \varepsilon_{i,k} + \pi_S(a_i - \lambda_{b_i,S}) I_{\rho(j)=S} + \pi_L(a_i - \lambda_{b_i,L}) I_{\rho(j)=L}.$$

Conditional on period-one city and birth state, the expected mean of $u_{i,j,k}$ is zero. Therefore, I can estimate g_j, π_S and π_L by running a regression in accordance with (11).

3. Empirical implementation

3.1. Census data

The primary data source is the Census Public Use Microdata 5 Percent Sample from 2000 (Ruggles et al., 2010). I focus on a sample of white men, aged 22 to 31 as of December 31, 1999, whose highest education attainment is a bachelor's degree. For these individuals, the Census reports the state of birth, the MSA of residence 5 years prior to the survey (i.e., residence in 1995), and the MSA of current residence.

I limit my analysis to individuals who reported working at least 40 weeks, at least 35 hours per week and earning at least 75 percent of the federal minimum wage in 1999.[3] I focus on college graduates with no graduate degrees because the timing of their labor-force entry varies less than that of other education groups. I exclude individuals in the primary sector, those with military history, and those born outside of the contiguous U.S. Table 1 presents the summary statistics for these two age cohorts.

My sample covers 328 localities of labor-force entry. Among them, 281 are MSAs, and the remaining 47 each represent the rural area of a state. The MSAs having more than 1.5 million people in 2000 are classified into the large-city category. Rural areas and MSAs with less than 1.5 million people are classified into the small-city category.[4] Fig. 1 presents the 36 large cities in my sample. Table 1 shows that nearly 50 percent of the college graduates in my sample entered the labor force in a large city.

3.2. Wage growth measure

The age groups 22–26 and 27–31 are adopted as the empirical counterparts for young and old workers in the model in order to construct a measure of wage growth. For the younger cohort, the current residence is assumed to be the location in which these individuals entered the labor force. For the older cohort, their residence in 1995 is assumed to be the location in which they entered the labor force.

The model assumes that the younger and older cohorts in the same city k receive the same initial wage levels w_k (Eq. (1) and (2)).[5] I first estimate w_k by running the regression

$$w_{i,k}^Y = w_k + X_i \tilde{\varphi} + u_i^Y \tag{12}$$

with the sample of the younger cohort. In Eq. (12), $w_{i,k}^Y$ is the log of the hourly wage for individual i in city k; X_i is a vector of control variables including age, age squared, marital status, and industry dummies; and u_i^Y is an error term.[6] The wage growth rate of the older cohort is obtained by subtracting the estimated initial wage levels ($\hat{w}_k$) from their wages; i.e.,

$$\Delta w_{i,j,k} = w_{i,j,k}^O - \hat{w}_k, \tag{13}$$

where $w_{i,j,k}^O$ is the log of the hourly wage for individual i in city k who lived in city j in 1995.

3.3. Large-city fraction by birth state

Because the information on the location of labor-force entry of the younger cohort is more accurate than that of the older cohort, I use the younger cohort to calculate P_b, the fraction of college graduates born in state b who entered the labor force in a large city.[7] Fig. 2 shows that P_b ranges from 0.14 to 0.80. This large variation

Table 1
Summary statistics.

	27 to 31	22 to 26
Migration history		
Entered the labor force in large cities (%)	42.5	49.0
Switched cities after labor-force entry (%)	43.8	
Labor market information		
Weeks worked in 1999	51.3	50.9
	(2.2)	(2.7)
Usual hours worked per week in 1999	46.9	45.5
	(8.4)	(8.0)
Hourly wage rate	21.9	16.3
	(15.7)	(9.9)
Industry sectors		
Construction (%)	4.1	4.1
Manufacturing (%)	16.4	14.5
Transportation and warehousing (%)	2.6	2.0
Utilities (%)	2.8	2.4
Wholesale (%)	5.4	4.0
Retail trade (%)	11.3	11.0
Business services (%)	27.5	28.4
Personal services (%)	1.9	2.1
Entertainment and recreation services (%)	3.2	3.5
Professional and related (%)	16.8	20.2
Public administration (%)	7.9	7.8
Other demographics		
Age	28.9	24.4
	(1.4)	(1.3)
Single	0.4	0.7
Observations	43,034	35,063

Notes: Standard deviations are in parentheses. Large cities include only MSAs of more than 1.5 million people in the 2000 Census.

[3] The full-time, full-year limitation allows me to focus on individuals who are less likely to be constrained to staying in their residential locations by family or education considerations. Measurement error is an additional justification for using full-time, full-year workers because Baum-Snow and Neal (2009) demonstrate that significant measurement error exists in hourly wages for part-time and part-year workers in the Census data.

[4] In Section 4.4.3 I show that 1.5 million is a reasonable city-size cutoff point.

[5] I assume that the city-specific initial wage levels remain constant between 1995 and 2000. Chen and Rosenthal (2008) find that the correlation of the local productivity of U.S. cities between 1990 and 2000 is 93.21%.

[6] The log of the hourly wage is calculated as the logarithm of wage and salary income divided by the product of weeks worked and usual hours worked per week.

[7] The main regression results do not change significantly if I use the older cohort to calculate P_b.

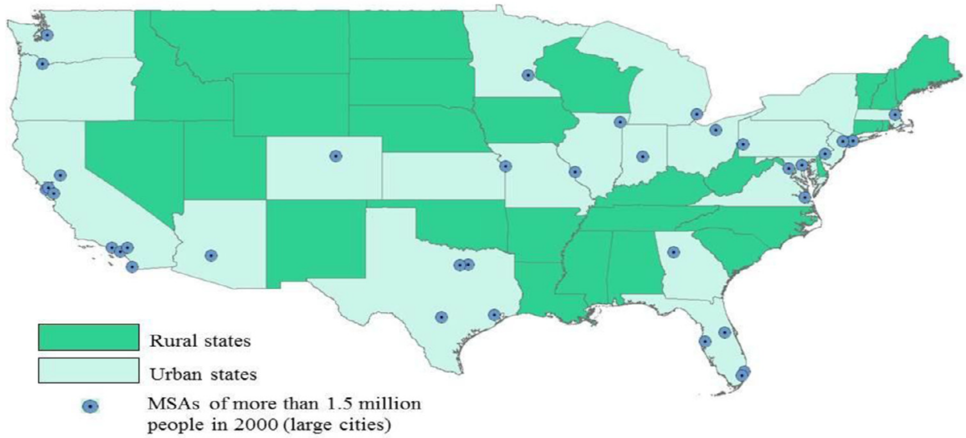

Fig. 1. Spatial distribution of large cities in the U.S.

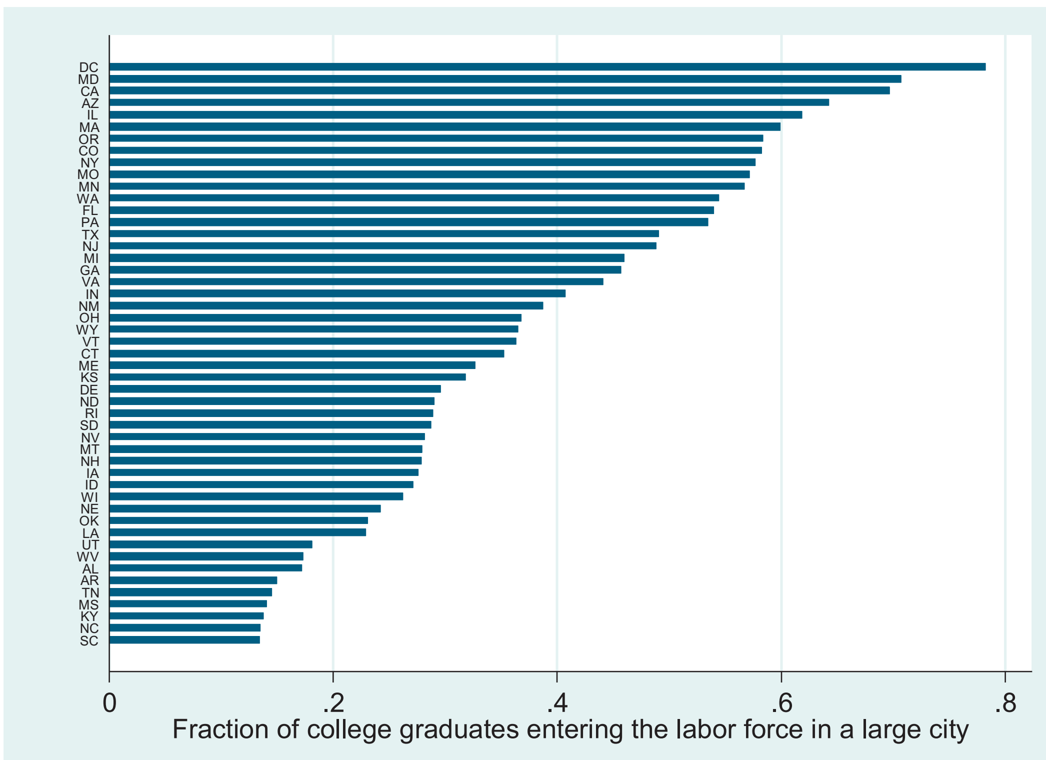

Fig. 2. Fraction of college graduates entering the labor force in a large city by birth state.

may be explained by the spatial pattern of large cities in the U.S. As shown in Fig. 1, its large cities are spatially concentrated. The dark-colored states represent states that have no large cities within their boundaries ("rural states"), while the light-colored states include at least one ("urban states").[8] According to the migration literature (e.g., Schwartz, 1973), people born in rural states who migrate to large cities may have to make more return trips to maintain social connections and, therefore, may need to bear a higher permanent transportation cost, which lowers the utility they gain from living in large cities. Hence, the selection threshold in (6) for rural states should be higher and the fraction of those who migrate to large cities should be smaller. Indeed, as shown in Fig. 2, college graduates born in rural states such as AR, TN, MS, KY, and SC were less likely to obtain their first jobs in large cities, while the majority of college graduates born in urban states such as DC, MD, CA, IL, and MA were more likely to enter the labor force in large cities.[9]

3.4. Specification of wage-growth regression

In my sample, the relationship between the ability proxies $\lambda_{b,\,S}$ and $\lambda_{b,\,L}$ in the wage-growth regression model (11) is almost lin-

[8] In the contiguous U.S., 22 of the 49 states (including Washington, DC and excluding Hawaii and Alaska) had an MSA of more than 1.5 million people in 2000 within their boundaries. An MSA's boundaries may cross state boundaries. The Philadelphia metropolitan area, for example, is shared between NJ and PA, and MO and IL share the St. Louis metropolitan area.

[9] Besides transportation costs, people born in different states may value large-city consumption amenities differently, thereby causing the selection threshold in (6) to vary across birth states.

ear, as both are approximately linear in P_b.[10] Therefore, for regression purposes, I apply the following linear approximations:

$$\lambda_{b,S} \approx -\tilde{\lambda}_S - \delta_S P_b, \tag{14}$$

$$\lambda_{b,L} \approx \tilde{\lambda}_L - \delta_L P_b. \tag{15}$$

I regress $\lambda_{b,\,S}$ and $\lambda_{b,\,L}$ on P_b, weighting each state by the number of individuals in the younger cohort born in that state. The estimates of $\tilde{\lambda}_S$, δ_S, $\tilde{\lambda}_L$, and δ_L are 0.02, 1.59, 1.75, and 1.84, respectively.

I assume that the wage growth specific to the locality of labor-force entry j in (11) can be expressed as a linear function of locality j's attributes Z_j:

$$g_j = Z_j \gamma. \tag{16}$$

Substituting (14)-(16) into (11), I obtain

$$\Delta w_{i,j,k} = \gamma_0 + Z_j\gamma + \alpha_1 I_{\rho(j)=L} + \alpha_2 I_{\rho(j)=L} P_{b_i} + \alpha_3 P_{b_i} + X_i\varphi + u_{i,j,k}, \tag{17}$$

where

$$\alpha_1 = \pi_L\tilde{\lambda}_L + \pi_S\tilde{\lambda}_S, \quad \alpha_2 = -\pi_L\delta_L + \pi_S\delta_S, \quad \text{and} \quad \alpha_3 = -\pi_S\delta_S.$$

In (17), γ_0 is a constant term, X_i includes age, age squared, marital status, and industry dummies, and $u_{i,\,j,\,k}$ is the error term. The parameters π_S and π_L are identified from the estimates of α_2 and α_3 in (17). For the main regression, Z_j includes locality j's population size, population density, and share of college graduates in the labor force.[11] All parameters in γ are expected to be positive because these three attributes are all associated with greater knowledge spillovers. (For a review, see Rosenthal and Strange, 2004.) Given that the data on Z_j are not available for the rural areas of the states, I include 47 rural dummies in the regression to capture the average wage growth of these localities.

The standard errors in the estimation are two-way clustered by the locality of residence in 1995 and by birth state. To address the concern that the estimated standard errors may be inconsistent because of the sampling error in the linear approximations in (14) and (15), I verify the robustness of the inference by examining the bootstrapped standard errors of the wage-growth regression. Details are presented in Section 4.1.

4. Main results from the Census data

4.1. Differentials in wage growth across U.S. cities

Table 2 displays the results of the wage-growth regression (17). Column 1 reports the results using the full sample of the 27–31 age cohort. The coefficients on the three locality attributes are all positive and statistically significant. On average, wage growth is faster if a worker enters the labor force in a larger, more densely populated city with a more highly educated population.

Because the rural area of a state may include multiple labor markets, the estimated initial wage levels may be less accurate for the workers in rural areas. Column 2 reports the results of the regression that excludes current rural residents. The results are little affected.

In the wage-growth regression, city-specific wage growth (g_j) is expressed as a linear function of three human-amenity attributes. I re-run the wage-growth regression with an additional set of locality attributes representing natural amenities and industrial composition.[12] Results not reported here show that the estimates of the coefficients on the three human-amenity attributes remain roughly unchanged, and the estimates of the coefficients on the other attributes are all statistically insignificant.

Table 3 presents the top and bottom 20 localities in the distribution of g_j, based on the coefficients from column 1 of Table 2.

Table 2
Wage-growth regressions.

Panel A: Wage-growth regression results		
Dependent variable: Wage growth, the 27–31 age cohort	1 Full sample	2 MSA sample
Log(population size)	0.020***	0.021***
	(0.004)	(0.005)
Population density	0.008***	0.008***
	(0.003)	(0.003)
Share college graduates in labor force	0.230***	0.211***
	(0.034)	(0.039)
Dummy: large city at labor-force entry	0.022	0.033*
	(0.018)	(0.020)
Dummy: large city at labor-force entry * Birth state's large-city fraction	−0.055*	−0.078**
	(0.031)	(0.034)
Birth state's large-city fraction	−0.025	−0.011
	(0.019)	(0.022)
R-squared	0.103	0.099
Observations	43,034	35,289
Panel B. Effects of ability on wage growth in large and small city categories		
Ability effect in large (π_L)	0.043***	0.049***
	(0.014)	(0.015)
Ability effect in small (π_S)	0.016	0.007
	(0.012)	(0.014)

Notes: Standard errors in parentheses are 2-way clustered by the locality of residence in 1995 and by birth state. All columns include age, age squared, marital status, 10 industry dummies, and a constant term.
* significance at 10%.
** significance at 5%.
*** significance at 1%.

Table 3
Localities with highest and lowest wage growth.

Rank	MSA/non-MSA Name	Rank	MSA/non-MSA Name
1	Jersey City, NJ	309	Jacksonville, NC
2	New York-Northeastern NJ	310	Rural, VA
3	Boston, MA	311	Rural, TN
4	Washington, DC/MD/VA	321	Houma-Thibodoux, LA
5	San Francisco-Oakland-Vallejo, CA	313	Rural, WI
6	Chicago-Gary-Lake, IL	314	Danville, VA
7	Los Angeles-Long Beach, CA	315	Rural, MN
8	Orange County, CA	316	Rural, PA
9	Nassau Co, NY	317	Rural, KS
10	San Jose, CA	318	Rural, MT
11	Philadelphia, PA/NJ	319	Rural, AR
12	Bergen-Passaic, NJ	320	Rural, NM
13	Oakland, CA	321	Rural, MO
14	Stamford, CT	322	Rural, NY
15	Atlanta, GA	323	Rural, NC
16	Middlesex-Somerset-Hunterdon, NJ	324	Rural, NV
17	Newark, NJ	325	Rural, LA
18	Seattle-Everett, WA	326	Rural, CA
19	Minneapolis-St. Paul, MN	327	Rural, MA
20	Baltimore, MD	328	Rural, AZ

[10] These approximately linear relationships still hold if I assume that ability follows other common distributions (e.g., exponential distribution, uniform distribution, and Chi-square distribution).

[11] The population size and the share of college graduates in the labor force of each of the 281 MSAs are calculated from the 2000 Census data, and the population density of each MSA is derived by dividing the population size by the land area provided by the U.S. Census Bureau.

[12] The natural-amenity attributes include the distance to the nearest body of water, annual precipitation, heating-degree days, and cooling-degree days. The industry-mix attributes include the employment share of business service, employment share of manufacturing, and ratio of business service employment to manufacturing employment.

New York, Boston, Washington, DC, San Francisco, Chicago, Los Angeles, and San Jose are among the top ten in the list of cities having the highest wage growth. Localities ranked toward the bottom are mainly rural areas.

For an average-ability worker ($a_i = 0$) who entered the labor force in a large city, the initial wage premium upon labor-force entry is 15.7 percent, and the wage-growth premium after 5 years is 7.2 percent.[13] Therefore, the urban wage premium for an average-ability worker 5 years after labor-force entry is 22.9 percent.

The standard errors in the regressions are estimated by two-way clustering the residuals on the locality of residence in 1995 and on the birth state. The results of all tests would remain unchanged if the residuals were clustered by either the locality of residence in 1995 or by the birth state, or if the standard errors were calculated on the basis of 1000 bootstrap replications clustering on the locality of residence in 1995 and on the birth state.

4.2. Effects of ability on wage growth in cities of different sizes

Panel B of Table 2 reports the estimates of π_S and π_L. The effect of ability on wage growth in large cities is positive and significant. In column 1, the estimate of π_L from the full sample suggests that a one-standard-deviation increase in ability raises wage growth by 4.3 percent in large cities, which is about 60 percent of the wage-growth premium for an average-ability worker entering the labor force in a large city. By contrast, the estimates of π_S in both columns 1 and 2 are small and statistically insignificant.

These results support the hypothesis that the wage-growth advantage of larger cities increases with ability (i.e., $\pi_L > \pi_S$), which causes higher-ability college graduates to sort themselves into larger cities. The same sorting pattern may also be due to high-ability workers' greater preference for large-city consumption amenities (Lee, 2010). However, the consumption-side explanation would imply that the urban wage premium should decrease with ability, which is contrary to the findings here.

4.3. Importance of ability sorting in the urban wage premium

The average wage-growth gap between large and small cities due to ability sorting can be estimated as follows:

$$\sum_b \tau_b^L \hat{\pi}_L \lambda_{b,L} - \sum_b \tau_b^S \hat{\pi}_S \lambda_{b,S}, \tag{18}$$

where $\lambda_{b,\,L}$ and $\lambda_{b,\,S}$ are proxies for the average ability of workers born in state b who entered the labor force in large and small cities, respectively, and $\hat{\pi}_L$ and $\hat{\pi}_S$ are defined as the estimates of the effects of ability on wage growth in large and small cities. In (18), τ_b^L and τ_b^S represent birth state b's share of college graduates who enter the labor force in large and small cities, respectively. These shares are based on the locations of labor-force entry of the 27–31 age cohort. The calculation of (18) indicates that, besides the urban wage premium of average-ability workers of 22.9 percent, the sorting of high-ability workers into large cities leads to another 3.2 percent of the urban wage-growth premium, suggesting that the selection accounts for around 10 percent of the urban wage premium.

Table 4
Wage-growth regressions with additional birth-state controls.

Panel A: Wage-growth regression results			
Dependent variable: Wage growth, the 27–31 age cohort	1 Full sample	2 Full sample	3 Full sample
Log(population size)	0.020***	0.020***	0.020***
	(0.004)	(0.004)	(0.004)
Population density	0.008***	0.008***	0.008***
	(0.003)	(0.003)	(0.003)
Share of college graduates in labor force	0.230***	0.230***	0.216***
	(0.034)	(0.037)	(0.035)
Dummy: large city at labor-force entry	Y	Y	Y
Dummy: large city at labor-force entry * Birth state's large-city fraction	Y	Y	Y
Birth state's large-city fraction	Y	Y	Y
Dummy: large city at labor-force entry * Birth state's share of highly-educated parents		Y	
Birth state's share of highly-educated parents		Y	
Dummy: large city at labor-force entry * Birth state's average reading score			Y
Birth state's average reading score			Y
Observations	43,034	43,034	43,034
Panel B. Effects of ability on wage growth in large and small city categories			
Ability effect in large (π_L)	0.043***	0.045***	0.047***
	(0.014)	(0.017)	(0.013)
Ability effect in small (π_S)	0.016	0.012	0.019
	(0.012)	(0.014)	(0.012)

Notes: Standard errors in parentheses are 2-way clustered by the locality of residence in 1995 and by birth state. All columns include age, age squared, marital status, 10 industry dummies, and a constant term. The share of highly-educated parents is measured with the share of college graduates among married males aged between 40 and 60 in 2000 in each state.
* significance at 10%.
*** significance at 1%.

4.4. Specification checks

4.4.1. Different distributions of ability across birth states

I assume in the baseline model that college graduates born in different birth states have identical distributions of ability. In reality, high-ability parents may have sorted themselves into urban states when they were younger. As a result, college graduates born in urban states, on average, may have higher ability than those born in rural states. Differences in infrastructure and school quality may also lead to different ability distributions across birth states.[14]

To address this issue, I allow m_b, the average ability of birth state b, to vary with birth state and re-run the wage-growth regressions using each state's average reading score and the share of highly-educated parents as proxies for m_b. The regression specification is

$$\begin{aligned}\Delta w_{i,j,k} &= \gamma_0 + Z_j\gamma + \alpha_1 I_{\rho(j)=L} + \alpha_2 I_{\rho(j)=L} P_{b_i} + \alpha_3 P_{b_i} \\ &\quad + (\pi_L - \pi_S) I_{\rho(j)=L} m_{b_i} + \pi_S m_{b_i} + X_i\varphi + u_{i,j,k}.\end{aligned} \tag{19}$$

[13] The initial wage premium upon labor-force entry and the wage-growth premium after 5 years for an average-ability worker are calculated by

$$\sum_{\{j:\rho(j)=L\}} \theta_j^L \hat{w}_j - \sum_{\{j:\rho(j)=S\}} \theta_j^S \hat{w}_j\,,$$

and

$$\sum_{\{j:\rho(j)=L\}} \theta_j^L \hat{g}_j - \sum_{\{j:\rho(j)=S\}} \theta_j^S \hat{g}_j\,,$$

where $\hat{w}_j$ and $\hat{g}_j$ are the estimates of the initial wage level and wage growth of city j, respectively, θ_j^L is large city j's share of college graduates who enter the labor force in a large city, and θ_j^S is small city j's share of college graduates who enter the labor force in a small city. These shares are based on the labor-force entry locations of the 27-31 age cohort.

[14] Census data show that parents in urban states have greater educational attainment than those in rural states. Additionally, the National Assessment of Education Progress (NAEP) shows that urban states (e.g., MA, MD, NJ, and TX) have higher average high-school test scores than rural states (e.g., WV, MS, AL, AR, TN, SC, and NC).

Table 5
Wage-growth regressions for the movers and stayers.

Dependent variable: Wage growth, the 27–31 age cohort	1 Full sample	2 Movers	3 Stayers	4 t stat (2-3)
Log(population size)	0.020***	0.025***	0.015*	0.894
	(0.004)	(0.006)	(0.008)	
Population density	0.008***	0.008***	0.008*	0.112
	(0.003)	(0.003)	(0.005)	
Share of college graduates in labor force	0.230***	0.239***	0.216***	0.285
	(0.034)	(0.060)	(0.054)	
Dummy: large city at labor-force entry	0.022	0.020	0.025	−0.148
	(0.018)	(0.032)	(0.017)	
Dummy: large city at labor-force entry * Birth state's large-city fraction	−0.055*	−0.043	−0.059*	0.275
	(0.031)	(0.051)	(0.031)	
Birth state's large-city fraction	−0.025	−0.019	−0.028	0.242
	(0.019)	(0.026)	(0.027)	
R-squared	0.103	0.122	0.091	
Observations	43,034	18,615	24,419	

Notes: Standard errors in parentheses are 2-way clustered by the locality of residence in 1995 and by birth state. All columns include age, age squared, marital status, 10 industry dummies, and a constant term.
* significance at 10%.
*** significance at 1%.

In (19), m_b affects wage growth in large cities as $(\pi_L - \pi_S)$ is positive. Because m_b is positively correlated with P_b, omitting $I_{\rho(j)=L}m_b$ from the wage-growth specification should lead to an overestimation of the coefficient on $I_{\rho(j)=L}P_b$ and an underestimation of π_L. As shown in Table 4, the inclusion of m_b leads to a moderate increase in the magnitude of π_L but the main results do not change.[15]

4.4.2. Movers and stayers

In my sample, about 44 percent of individuals in the 27–31 age cohort moved across MSA boundaries during the 5-year period after labor-force entry. These individuals are referred to as movers. The model implies identical coefficients between the movers and stayers from the wage-growth regression.[16] I re-run the regressions separately for the movers and stayers. The results are reported in columns 2 and 3 of Table 5. (Column 1 replicates the results from column 1 of Table 2.) Column 4 of Table 5 presents the t statistics for testing the equality of each coefficient in columns 2 and 3. The results indicate that the estimates for the movers are not statistically different from those of the stayers.[17]

4.4.3. Other city-size cutoff points

The choice of the cutoff point for defining a large city is crucial to the estimation of π_S and π_L. In the main analysis, I use 1.5 million as the cutoff point. Lowering the cutoff point increases the number of cities in the large-city category and reduces the gap between the ability effects on wage growth in large and small cities. Fig. 3 plots the t statistics for the estimated $(\pi_L - \pi_S)$ for different city-size cutoff points. Note that the t statistic peaks at 2.12 at the 1.5 million cutoff point, suggesting that the city-size cutoff used in this paper is reasonable.

5. Corroborating evidence from the NLSY79

This section provides corroborating evidence from the NLSY79 data, which contain individual AFQT scores that measure individual ability before labor-force entry.[18] My sample is drawn from a confidential version of the NLSY79 data that provides information on the MSA of residence of each respondent in each survey round from 1979 through 1994. It contains 1974 observations of 258 white men whose highest educational attainment as of 1994 was a bachelor's degree.[19] Nearly 90 percent of the sample had over six years of work experience. An MSA with a population above 1.2 million in the 1980 Census is defined as a large city.[20] In the sample, 144 individuals entered the labor force in a large city.[21]

Fig. 4 graphs the bivariate relations between the log of the real hourly wage and years of work experience for individuals who entered the labor force in small cities (left panel) and large cities (right panel).[22] In each panel, the solid line represents the average wage-experience profile for individuals with above-median AFQT scores, while the dashed line represents the average wage-experience profile for individuals with below-median AFQT scores. Note that the wage-experience profile is steeper in large cities than in small cities. Furthermore, the wage-experience profiles of high- and low-AFQT individuals diverge only in large cities. These patterns are consistent with the results shown in Sections 4.1 and 4.2.

To check the robustness of Fig. 4, I estimate the model

$$w_{i,t} = \beta_0 e_{i,t} + \beta_1 e_{i,t} \times I_i + \tilde{\pi}_S(AFQT_i \times e_{i,t}) + \tilde{\pi}_L(AFQT_i \times e_{i,t} \times I_i) + \Omega_{i,t}\vartheta + \xi_{i,t}. \tag{20}$$

In (20), $w_{i,\,t}$ is the log of the real hourly wage of individual i in year t, I_i is a dummy variable that equals 1 if individual i enters the labor force in a large city, $e_{i,\,t}$ is the experience level in year t, $AFQT_i$ is the AFQT score of individual i, and $\xi_{i,\,t}$ is an error term. The vector $\Omega_{i,\,t}$ includes an indicator for urban residence in year t, $AFQT_i$, year fixed effects, the square and cubic terms of $e_{i,\,t}$, and a constant term.[23] The coefficient β_1 captures the wage-growth premium of entering the labor force in a large city, and the coefficients $\tilde{\pi}_S$ and $\tilde{\pi}_L$ represent the effects of ability on wage growth in small and large cities, respectively. Column 1 of Table 6 presents

[15] The NAEP provides state-level average scores for public school students within a single assessment year. I document the average reading, math, writing, and science scores of white students in grade 8 in 2011 for the 49 states. Results using math, writing, and science scores not reported here are similar to those using reading scores. The share of highly educated parents is measured with the share of college graduates among married males aged between 40 and 60 in 2000 in each state.

[16] In the model, wage growth is portable when a worker moves to another city, and the move after labor-force entry is independent of individual ability.

[17] The F statistic for simultaneously testing the equality of all coefficients is 0.64 and the p value is 0.70.

[18] The NLSY79 is a nationally representative sample of men and women who were between the ages of 14 and 22 when they were selected at the beginning of 1979. The AFQT, or Armed Forces Qualification Test, consists of the following four sections from the Armed Services Vocational Aptitudes Battery: word knowledge, paragraph comprehension, arithmetic reasoning, and mathematics knowledge.

[19] The original sample includes 494 white men who received their bachelor's degrees prior to 1994 from the NLSY79 random sample of 3,003 men. 46 individuals are dropped because they were in the military at some point. An additional 10 individuals are dropped because they received their bachelor's degrees prior to 1979. Only jobs worked after an individual has left school for the first time are kept in the sample. I keep individuals who entered the labor force prior to age 30. 90% of them entered the labor market between the ages of 22 and 26. Individuals with missing wage information at the time of labor-force entry, individuals whose wages were below $1 or above $100, and individuals with missing AFQT scores are excluded. Finally, I drop individuals whose labor-force entry locations are unidentifiable. The resulting sample has 1,974 observations of 258 individuals.

[20] The lists of the MSAs having more than 1.2 million people in 1980 and of those having more than 1.5 million people in 2000 largely overlap.

[21] An individual whose first or second observation appeared in an MSA with a population above 1.2 million in the 1980 Census is treated as entering the labor force in a large city.

[22] Real wages in 1984 prices are calculated using the CPI deflators from the U.S. Census Bureau.

[23] The AFQT scores are standardized to have a mean of zero and standard deviation of one.

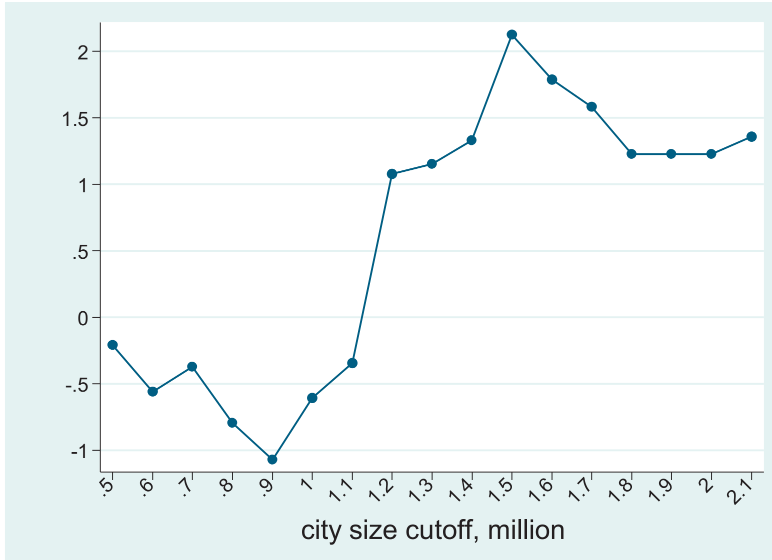

Fig. 3. T statistics for estimated differences between the ability effects on wage growth in large and small cities for different city-size cutoff points.
Notes: Each dot represents the t statistic for the estimate of the gap between the ability effects on wage growth in large and small cities. the estimates are obtained from wage-growth regression (17).

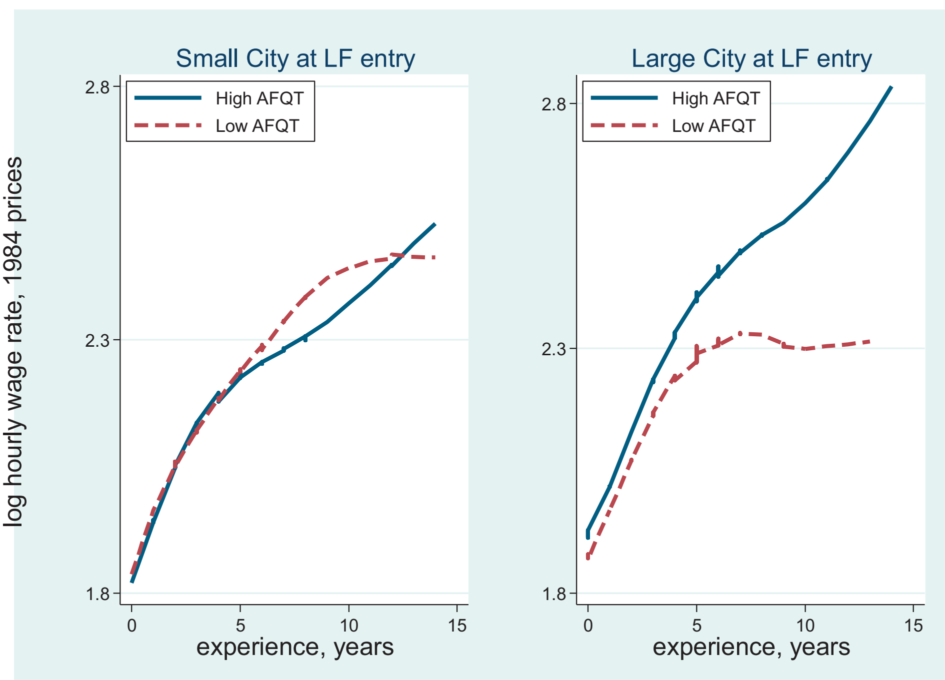

Fig. 4. Wage-experience profiles, the NLSY79.
Notes: The plot comes from the estimation of nonparametric regressions that use a uniform kernel density regression smoother.

the regression results for the full sample. Consistent with the findings from the Census data, the estimates of β_1 and $\tilde{\pi}_L$ are positive and statistically significant, while the estimate of $\tilde{\pi}_S$ is small and statistically insignificant.

Columns 2 and 3 of Table 6 report the results for the subsamples of movers and stayers, respectively.[24] As shown in column 4 of Table 6, the t statistics for testing the equality of each coefficient in columns 2 and 3 suggest that the estimates are not statistically different between the movers and stayers.[25] The results are consistent with the results in Section 4.4.2.

6. Conclusion

This paper proposes a novel method of identifying selective migration in the estimation of the effects of cities on wage growth using the relatively limited mobility information contained in the Census data. The results suggest that a young worker's labor-force

[24] An individual is classified as a mover if he changed his MSA of residence at least once after entering the labor force.

[25] The F statistic for simultaneously testing the equality of all coefficients is 0.14 and the p value is 0.97.

Table 6
Wage-experience regressions, the NLSY79.

Dependent variable: log(real hourly wage rate)				
	1 Full sample	2 Movers	3 Stayers	4 t stat (2-3)
Experience	0.133***	0.126***	0.151***	−0.527
	(0.024)	(0.030)	(0.037)	
Experience* Dummy: large city at labor-force entry	0.012***	0.012***	0.011*	0.102
	(0.004)	(0.004)	(0.006)	
Experience* AFQT	−0.004	−0.005	−0.002	−0.400
	(0.004)	(0.005)	(0.007)	
Experience* Dummy: large city at labor-force entry* AFQT	0.013***	0.014***	0.010*	0.507
	(0.004)	(0.004)	(0.006)	
R-squared	0.175	0.151	0.272	
Observations	1974	1417	557	

Notes: Heteroskedasticity-robust standard errors are in parentheses. All columns include the individual's AFQT score, an indicator for urban residence in year t, year fixed effects, the square and cubic terms of experience, and a constant term.
* significance at 10%.
*** significance at 1%

entry location affects his subsequent wage growth. Furthermore, individual ability boosts wage growth in large cities but not in small cities or rural areas. This causes the selection of higher-ability workers into larger cities at labor-force entry. I also provide corroborating evidence from the NLSY79 data. Further research could apply this method to investigate other demographic groups. In addition, the method could be used to analyze Census data from other years in order to understand the evolution of the selection effects, wage-growth effects, and wage-level effects of U.S. cities.

Acknowledgments

I am grateful to Vernon Henderson, Nathaniel Baum-Snow, and Andrew Foster for advising me on this project. I am greatly indebted to the editor, two anonymous referees, and Jimmy Chan, who helped improve the paper. I also thank Anna Aizer, Ken Chay, Blaise Melly, Sriniketh Nagavarapu, Anja Sautmann, Michael Suher, and seminar participants at Brown University, the Urban Economics Association 2012 Annual Meeting in Ottawa, and the Econometric Society Asian Meeting 2013 for helpful comments and suggestions. Any remaining errors are mine.

References

Chen, Y., Rosenthal, S.S., 2008. Local amenity and life-cycle migration: do people move for jobs or fun? J. Urban Econ. 64, 519–537.
Baum-Snow, N., Pavan, R., 2012. Understanding the city size wage gap. Rev. Econ. Stud. 79 (1), 88–127.
Baum-Snow, N., Neal, D., 2009. Mismeasurement of usual hours worked in the Census and ACS. Econ. Lett. 102 (1), 39–41.
Duranton, G., Puga, D., 2004. Micro-foundations of urban agglomeration economies. In: Henderson, J.V., Thisse, J.F. (Eds.). In: Handbook of Regional and Urban Economics, vol. 4. North-Holland, Amsterdam, pp. 2063–2117.
Glaeser, E.L., 1999. Learning in cities. J. Urban Econ. 46, 254–277.
Glaeser, E.L., Mare, D.C., 2001. Cities and skills. J. Labor Econ. 19 (2), 316–342.
Gould, E.D., 2007. Cities, workers, and wages: a structural analysis of the urban wage premium. Rev. Econ. Stud. 74, 477–506.
Lee, S., 2010. Ability sorting and consumer city. J. Urban Econ. 68, 20–33.
Roback, J., 1982. Wages, rents, and the quality of life. J. Polit. Econ. 90 (6), 1257–1278.
Rosenthal, S.S., Strange, W.C., 2004. Evidence on the nature and sources of agglomeration economies. In: Henderson, J.V., Thisse, J.F. (Eds.). In: Handbook of Regional and Urban Economics, vol. 4. Elsevier, New York, pp. 2119–2171.
Ruggles, S., Alexander, T.J., Genadek, K., Goeken, R., Schroeder, M.B., Sobek, M., 2010. Integrated Public Use Microdata Series: Version 5.0. University of Minnesota, Minneapolis [Machine-readable database].
Schwartz, A., 1973. Interpreting the effect of distance on migration. J. Polit. Econ. 81 (5), 1153–1169.
Wheeler, C.H., 2006. Cities and the growth of wages among young workers: evidence from the NLSY. J. Urban Econ. 60, 162–184.
Yankow, J.J., 2006. Why do cities pay more? an empirical examination of some competing theories of the urban wage premium. J. Urban Econ. 60, 139–161.