Applied Spatial Analysis and Policy (2020) 13:631–648
https://doi.org/10.1007/s12061-019-09322-6

China's Highly Educated Talents in 2015: Patterns, Determinants and Spatial Spillover Effects

Hengyu Gu[1,2] · Xin Meng[1] · Tiyan Shen[1] · Luge Wen[1]

Received: 21 June 2019 / Accepted: 27 August 2019 /
Published online: 12 October 2019

Abstract
Using data from the 2015 national one percentage population sample survey, this paper examines the distribution, driving forces, and spatial effect of highly educated talents at the prefecture level. It involves spatial autocorrelation analysis and spatial econometrics models. The results show that the spatial pattern of talents is highly concentrated, unbalanced, and clustered among cities. Economic opportunities are the main forces affecting the distribution of talents, although some amenity variables (e.g., public services and accessibility) also matter. Our findings suggest that spatial spillover effect of the distribution of talents mainly comes from the influence of cross-city or country specific talent policies and social network linkages.

Keywords Highly educated talents · Spatial pattern · Driving forces · Spatial spillover effect · China

Introduction

Knowledge accumulation and technological progress are sources of regional sustained economic growth (Lucas 1988; Romer 1990). As carriers of knowledge and technology, highly educated talents are key drivers of national and regional development. How to attract talents, retain talents, and make use of talents effectively is a common concern of governments around the world (Lee et al. 2004). The concentration of talents can bring about a variety of agglomeration economies, such as knowledge spillovers, low-cost communication and learning externalities, and a specialized talent pool for certain industries that need intelligence (Fujita and Thisse 1996). China has experienced dramatic economic restructuring and industrial upgrading in recent years, and technological

✉ Tiyan Shen
tyshen@pku.edu.cn

[1] School of Government, Peking University, No.5 Yiheyuan Road, Beijing 100871, China

[2] Center for Spatial Data Science, University of Chicago, 5735 S Ellis Ave, Chicago, IL, USA

innovation has therefore become the core driving force for national and regional development (Duan et al. 2016; Liu and Shen 2014; Xu and Ouyang 2018). Talents are undoubtedly the first element to promote technological innovation and regional economic and social development (Florida et al. 2012). Therefore, Chinese central government has implemented several policies to illustrate the importance of highly educated talents to China's regional economic growth in the transition period. For example, in 2010, the State Council of China promulgated the first medium- and long-term talent-development planning document, namely *National Medium- and Long-Term Talent Development Plan (2010–2020).* Its outline points out that China must change from a country with huge human resources to a country with a huge amount of talents, and win an active strategic battle in the fierce international competition. Furthermore, in recent years, Chinese local governments have formulated a series of public policies for attracting and retaining highly educated talents (Yuan and Wan 2007), which have set off an intensified wave of competition for talents.

In recent years, the increase in China's highly educated talents has become an ongoing trend. According to the 1% national sample survey of the population in 2015, China has a population with a college education or above reaching 170.93 million, an increase of 42.87% over the same period in 2010. A burgeoning body of literature has focused on the spatial patterns of highly educated people at provincial or prefecture level in China and drivers of those patterns (Bao et al. 2007; Gu et al. 1999; Liu and Shen 2014; Liu et al. 2017; Miu et al. 2013; Wang and Zhang 2003; Zhang et al. 2011). According to these studies, there is an unbalanced trend of talent distribution in China, which has led to a regional imbalance of talent demand (Liu and Shen 2014; Liu et al. 2017). The concentration effect of talents in those developed cities has further squeezed the development room for undeveloped small and medium-sized cities. Furthermore, there is likely to be a positive spatial spillover effect in the distribution of talents across space because of a positive externality in the number of talents or a positive spatial autocorrelation in determinants associated to the number of talents in neighboring cities (Miguélez et al. 2010). However, few researches have been conducted on the spatial spillover effect in Chinese highly educated talents at city level. In addition, due to data availability, most studies focus on the distribution and determinants of all kinds of China's talents before 2010 (e.g., Liu and Shen 2014; Liu et al. 2017). Under this context, we hope to contribute to previous studies in following aspects: (a) using data from 1% national population survey, this paper examines the distribution and drivers of highly educated talents at city level for the first time; (b) using spatial econometrics methods, this study attempts to investigate whether the distribution of highly educated talents between cities shows a spatial spillover effect.

The main object of this paper is to examine the spatial pattern, determinants and spatial spillover effect of highly educated talents at prefecture level units in China, by using data derived from the 2015 national 1% population sampling survey data. The next section begins with a systematic summary of the spatial pattern and influencing factors of highly educated talents. The third section explains the research design, including the research data, regions and research methods. Then, there is a description of several GIS spatial analysis methods involved in the spatial analysis of the distribution of talents. The fifth section constructs a series of econometric models (including nonspatial models and spatial econometric models) to explore the impacts of the concentration of highly educated talents. Then, there are some conclusions, and the final section offers relevant talent management recommendations.

Literature Review

Since the pioneering of population economic theory, many have seen economic opportunities as the primary conditions affecting the migration of talents (e.g., Lewis 1954; Todaro 1969). Regional economic opportunities include specific elements, such as wages relating to the human capital of the talents themselves, employment environment conditions, the openness of job market, opportunities for talent future development, and regional industrial structure. In the middle of the twentieth century, economists, demographers, and geographers explored the mechanism behind labor mobility from the perspectives of macroeconomics and microeconomic individuals, based on hypothesis of rational-economic man, and they proposed a series of classical theories. From the microeconomic perspective, economists believed that each worker has sufficient economic rationality, and that the motivation for migration comes from economic benefits (Lewis 1954; Stark and Bloom 1985; Todaro 1969). When wage income is greater than migration cost, labor has sufficient motivation to move (Esenwein-Rothe 1964). Todaro (1969) set out the migration choices of laborers from a microscopic perspective. The motivation for labor migration lies in expected income rather than absolute income. The labor migration decision is the result of comparing expected income with migration costs. In the 1980s, the focus of migration economics theory shifted from the individual to the family. Stark and Bloom (1985) argued that labor migration is the result of household income diversification and risk minimization decisions. Labor migration occurs not only to maximize individual economic benefits, but also to achieve household economic income goals. Once laborers achieve their expected income targets, they may choose to settle down.

Krugman (1991) proposed the concept of new economic geography. In the framework of monopolistic competition, he put the transportation cost and manufacturing scale return increment into the model, and he proposed a core-periphery model (Krugman 1991). Regardless of externalities, the model can explain whether the spatial pattern of labor force is agglomerated or discrete simply by looking at increasing returns of scale and improved transportation costs (Fujita and Thisse 1996; Krugman 1991).

Another group of scholars began to study the impact of amenity factors on the migration of talents (e.g., Glaeser et al. 2001). Based on the perspective of equilibrium theory, Graves (1976) optimized the migration model of Cebula and Vedder (1973). Assuming that individuals' utilities in every region are the same in equilibrium, Graves (1976, 1979) found that the economic income of each region compensated for its level of amenities for laborers. Therefore, the migration of laborers is essentially the result of the need for regional amenities rather than economic income (Knapp and Gravest 1989; Mueser and Graves 1995). On the basis of Graves's (1976) equilibrium theory, Glaeser et al. (2001) put forward the hypothesis of the *consumer city*. They believed that the key to urban sustainable development lies in the quality of services and consumer goods available in the city. A city with a higher degree of amenities will therefore attract more talents to live in it, thus placing its economic development and innovation ahead of cities with low amenity levels.

Florida (2002) defined the conception of creative class, which transcended the definition of traditional highly skilled labors with academic or occupational divisions. He found that the reason for the accumulation of innovative talents was a city's facilities, including both its natural conditions, such as urban climate, and the inclusive humanistic environment of the city. Clark et al. (2002) pointed out that improvements in transportation technologies have reduced the costs of transporting goods and migrating, and a multi-center urban pattern has gradually formed. At this stage, the key to attracting talents lies in the supply of public goods and the physical facilities of the city (Clark et al. 2002).

Despite the increased focus on amenity factors in recent years, scholars have not reached a consensus on the incidences of economic opportunities and amenity factors, and therefore they have launched a heated discussion around this topic (Arntz 2010; Faggian et al. 2017; Liu and Shen 2014; Niedomysl et al. 2010; Scott 2010). Although several scholars acknowledged the importance of amenity factors, these factors should be secondary considerations in labor migration decision-making (Arntz 2010; Greenwood and Hunt 1989; Niedomysl and Hansen 2010). Other scholars believed that the impact of amenities, especially the natural environment, on the distribution of talents is still unverified (Liu and Shen 2014). For example, although several industrial cities in the interior of the United States do not have good living environments, they have attracted a large number of talents to work because they can provide a variety of employment opportunities (Chiquiar and Hanson 2005; Davies et al. 2001). In addition, the factors influencing talents gathering may be different at different historical stages, and the impact of job opportunities and local, quality-related factors varies among different life stages of the talents. There are also differences in the types of talents (e.g., high-educated immigrants and low-educated immigrants) (Chen and Rosenthal 2008; Ferguson et al. 2007; Niedomysl and Hansen 2010; Scott 2010). Gottlieb and Joseph (2006) found that compared with other academic talents, doctoral students pay less attention to wages when choosing employment locations, but instead they focus on amenity factors such as climate, public services, and crime rates.

Because of the spatial correlation of driving forces that affect talent distribution, such as economic linkages and policy response, there is also likely to be a spatial autocorrelation (i.e. spatial spillover) effect in the spatial distribution of talents. So far, a burgeoning but small body of literature has paid attention to the spatial spillover effect in the distribution of talents. Miguélez et al. (2010) have confirmed that there were a strong and positive spatial autocorrelation in the inflows of inventors across European regions, and that this spatial autocorrelation effect came from three major reasons: the attractiveness of a certain region, a certain region's congestion effect, and the influence of country specific features. Rodríguez-Pose and Tselios (2011) also found a positive spatial spillover effect in the distribution of education attainment in 102 regions in Western Europe. In recent years, scholars have begun to conduct several empirical studies using China as a case. First, research showed that economic opportunities are the primary force shaping the migration and distribution of talents in China (Gu et al. 2019a; Liu and Shen 2014; Yu et al. 2019). According to this stream of research, regional GDP, per capita income, unemployment rate, and industrial structure are the key forces influencing the spatial pattern of Chinese talents. Some scholars found that the western region in China was unable to attract talents because of its unsatisfactory talent market environment, talent entrepreneurial environment, and talent economic development environment (Bao et al. 2007). Some scholars started from the micro level, and they conducted research through questionnaires to determine the economic factors such as job opportunities, living conditions and wage levels, which are the main drivers of migration, further (Gu et al. 1999; Wang and Zhang 2003). In addition, scholars have further demonstrated the decisive role of economic opportunities through research on different types of talents, including highly skilled labor, scientific research talents, and highly educated talents (Miu et al. 2013; Zhang et al. 2011).

In addition, domestic scholars have begun to explore the role of amenity factors in the distribution of talents (e.g., Liu and Shen 2014). First, they found that public services such as health and education reflect the impacts of amenities. The attraction of talents in big cities is mainly the result of the supply of social security (Deng et al. 2001; Luo 2001; Zhang et al. 2017). Second, despite the importance of social and cultural factors, the impact of these variables on talents is not easy to observe through quantitative data; thus, a few scholars have used questionnaire methods to conduct research from the scope of micro-individuals (Li et al. 2018). Third, consumer facilities reflect the convenience of urban life, and it is the common psychology of talents to live in convenience and comfort (Liu et al. 2017). Fourth, because of the increasingly long commute time between cities, the impact of transportation facilities on talents is increasingly significant (Deng et al. 2001; Xu and Zhang 2010).

Furthermore, the relevant literature on comparing the impacts of economic opportunities and amenities agrees that economic opportunities are the main determinants shaping Chinese talents, whereas amenity variables, although significant in some studies (Song et al. 2016; Yu et al. 2019), are less attractive to talents (Liu and Shen 2014). However, existing research mainly uses the census data of 2000, 2005, or 2010. After the Chinese government proposed the *National Medium- and Long-Term Talent Development Plan (2010–2020)* in 2010, it is still unclear which forces actually affect the distribution of Chinese talents due to the lack of the necessary census data. In addition, the measurement of amenity in previous studies is not comprehensive. For example, in previous studies, researchers paid less attention to consumption diversity and regional traffic accessibility.

Also, a small body of studies have been conducted on the spatial spillover effect in the distribution of talents. A positive spatial spillover effect is also found in the spatial patterns of the birthplace and the workplace of academicians of Chinese Academy of Sciences over the period of 1955 to 2011 (Li et al. 2013). Using provincial level panel data, Song et al. (2016) has shown that a spatial autocorrelation effect was also seen in the employment of IT service sector and IT manufacturing sector. Research on spatial spillover effect is necessary for better understanding of the driving forces of talents' spatial moves or developing regional talent policies in China. However, there is little literature on spatial spillover effect in the distribution of highly educated talents. This paper hopes to address this literature gap by using spatial econometrics models.

Research Data, Area and Methodologies

Research Data

The basic data for our paper came from the 2015 national 1% population sample survey. From the dataset, there were 147,139 thousand samples with a college education or above in 2015, accounting for 10.71% of the total population. We define talents as "people with junior college degrees or above." The benefits of using academic qualifications to define talents are reflected in the fact that academic qualifications can better represent the human capital and labor capital of the labor force. In addition, relevant talent policies in Chinese cities often use academic qualifications to divide talents. Using academic qualifications to define talents can better match the actual talent needs of each city.

To ameliorate the endogeneity problem of reverse causality, the dependent variable in our article has a one-period lag from the independent variable. The independent variables came from the *China City Statistical Yearbook* from 2015, reflecting the economic and social conditions of cities at the end of 2014. The meteorological data for each city for 2014 came from the *National Weather Station of the National Meteorological Information Center*. We aggregated the weather data to local cities. The national road network (provincial road, national highway, highway) data necessary for calculating road network accessibility came from the *China Traffic Atlas* of 2015. Specifically, variables mainly fell into three sets: economic opportunities variables, amenity variables, and control variables.

Research Area

Our research area covers 31 provinces in China, excluding Hong Kong, Macao, and Taiwan. The basic geographic unit is the prefecture level, because the city is the basic unit for formulating talents and population management policies in China, and the differences in factors that influence the distribution of talents in various cities are huge. We aggregated highly educated talents with a college education or above to prefecture-level units according to the regional administrative code of each prefecture-level city, and we finally obtained 309 prefecture-level units as our basic units. We measured the spatial pattern of highly educated talents according to two types of statistic specifications: the number of highly educated talents and the density of highly educated talents (persons/km^2). In the regression model, considering the availability of data and the continuity of space, we selected a subset of 224 prefecture-level cities as the sample for regression analysis.

Research Methodologies

Gini Index (GI) The Gini coefficient measures the imbalance of highly educated talent distribution in 309 prefecture-level units in China, with its formula as follows (Liu and Gu 2019):

$$G = \frac{\sum_{i=1}^{n} lnx_i \times lny_i / \sum_{i=1}^{n} (lnx_i)^2 - 1}{\sum_{i=1}^{n} lnx_i \times lny_i / \sum_{i=1}^{n} (lnx_i)^2 + 1} \tag{1}$$

where x_i is the cumulative proportion of highly educated talents in each municipal unit until city i and y_i is the cumulative proportion of highly educated talents in each city until city i.

Exploratory Spatial Data Analysis Spatial autocorrelation of highly educated talents among cities is an aspect of understanding its spatial pattern. Spatial autocorrelation stems from the first law of geography, which means that the attribute values of objects have correlations with each other in neighboring regions (Tobler 1970). Derived from the Pearson correlation coefficient, Moran's I Coefficient (*MC*) is the most popular method of spatial autocorrelation (Gu et al. 2019b). The formula is:

$$MC = \frac{X'\mathbf{W}X}{X'X} \tag{2}$$

where X represents the column vector of observations and W represents the standardized spatial weight matrix. *MC* can measure the spatial autocorrelation of highly educated talent density in cities, and, to a certain extent, it represents whether the density or the number of highly educated talents in each city is agglomerated, randomly distributed, or discretely distributed.

Furthermore, we use local spatial autocorrelation (LISA) to measure the association of the density of highly educated talents in each city with its neighboring cities and to identify the spatial cluster pattern (Anselin 1995). The Formula is as follows:

$$Local\ Moran's\ I = \frac{n\left(x_i - \overline{x}\right)\sum_{j=1}^{m} W_{ij}\left(x_j - \overline{x}\right)}{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}, (i \neq j) \tag{3}$$

Where x_i and x_j are the density of talents of city i and city j; n is the number of cities; W_{ij} is the spatial weight matrix, and $\overline{x}$ is the average of x. Using LISA, four cluster patterns can be detected according to the association of talent density in one city and its neighboring cities. A High-High (HH) city has high talent density with average neighboring talent density above the mean. A Low-Low (LL) city has low talent density with average neighboring talent density below the mean. A High-Low (LH) city has high talent density with average neighboring talent density below the mean. A Low-High (LH) city has a low talent density with average neighboring talent density above the mean. Cities that fail statistical tests are indicated as not significant (NS).

Spatial Econometrics Models Spatial econometrics models deal with spatial effects in data by adding operators that reflect spatial autocorrelation in model settings (Stillwell et al. 2018). When using traditional linear or nonlinear models to examine economic drivers and amenity drivers of the spatial distribution of highly educated talents, it is easy to overlook the spatial autocorrelation problem in the data due to externality, unobservable variables, or missing variables. Therefore, the error term in the traditional regression does not satisfy the assumption of i.i.d., which leads to a larger estimation error or a decrease in the validity of the model. We employed a general spatial model (GSM), considering two spatial spillover effects (spatial lag and spatial error effects) together. The formula is as follows:

$$\mathrm{y} = \rho W_1 y + x\beta + \mu, \mu = \lambda W_2 \mu + \varepsilon, \varepsilon \sim \left(0, \sigma^2 I_n\right) \tag{4}$$

where $\rho W_1 y$ and $\lambda W_2 \mu$ represent spatial lag terms of the dependent variable and the random error term respectively, and ε is a random error term that satisfies the independent and identical distribution. Supported by these spatial econometrics, it is possible to examine the driving mechanism of talent distribution.

Spatial Pattern of Chinese Highly Educated Talents

A Spatially Concentrated and Unbalanced Pattern of Highly Educated Talents

Firstly, we calculated the GI of talent density (persons/km^2) of China's cities. The GI of talents reached 0.567 (> 0.5), indicating that the highly educated talents are very unevenly distributed among prefecture-level units. Secondly, we mapped the density of highly educated talents in each city using ArcGIS. As Fig. 1 shows, China's highly educated talents are highly concentrated in a small subset of cities. High-ranking cities include first-tier cities such as Beijing Shanghai, Guangzhou and Shenzhen, capital cities in the central and eastern regions such as Zhengzhou, Taiyuan, Wuhan, and Nanjing, and Dongguan, Zhuhai, and Zhongshan in the Pearl River Delta. In summary, the distribution of talent density of cities demonstrates a spatially concentrated and unbalanced pattern.

A Clustered Spatial Pattern of Highly Educated Talents

The global spatial autocorrelation analysis measures the overall level of spatial autocorrelation of the density and the number of highly educated talents among cities. Using Geoda software, it is possible to calculate the two spatial autocorrelation values. The *MC* value of the density of talents in China is 0.145 ($p < 0.001$), while this index for the number of highly educated talents reaches 0.073 ($p < 0.001$). The above results show that there is a significant spatial spillover effect in both the density and the number of high-level talents in prefecture-level cities. Furthermore, the LISA map can measure the similarity and dissimilarity of the attributes of particular geographic units to their surrounding units. In Geoda software, we obtained the LISA result of spatial clusters of Chinese talents. As Fig. 2 shows, the red parts are HH clusters, including several cities in the Yangtze River Delta urban agglomerations such as Shanghai and Hangzhou, and a few cities in the Pearl River

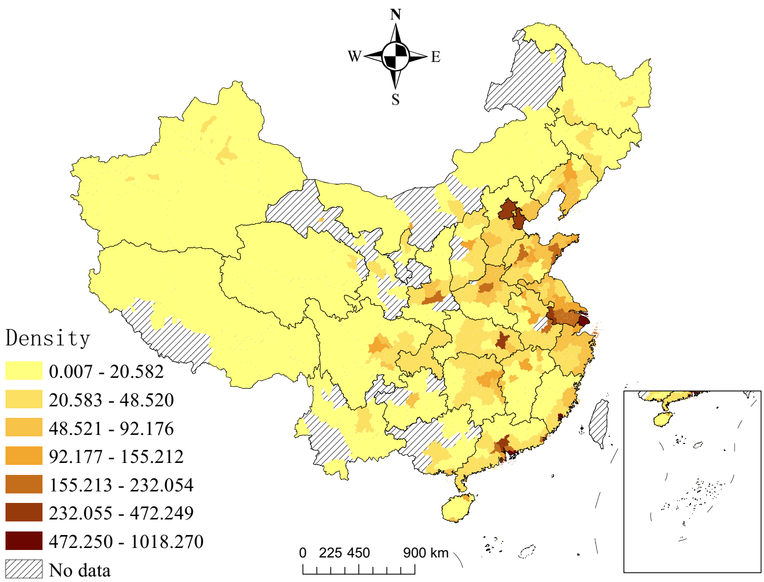

Fig. 1 Spatial distribution of density of highly educated talents in 2015

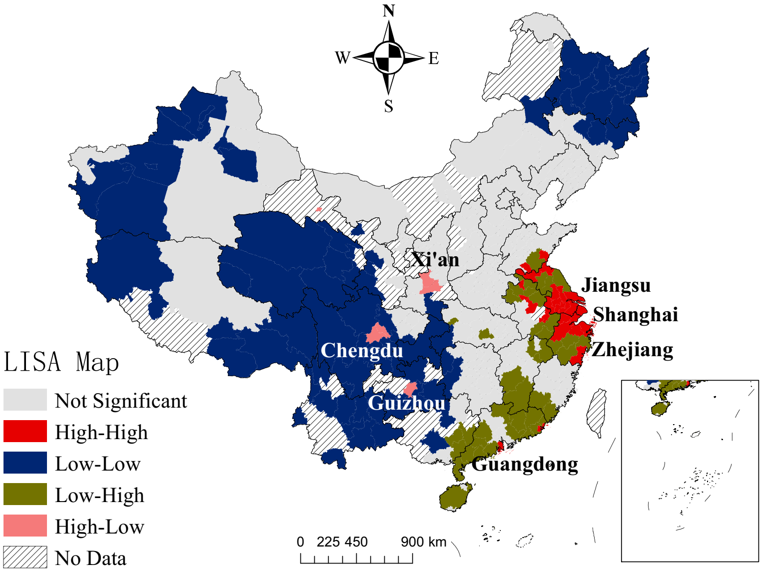

Fig. 2 LISA map of talent distribution pattern

Delta urban agglomerations like Zhongshan and Zhuhai. These two urban agglomerations have played a strong joint driving role in attracting talents. Some capital cities (e.g., Xi'an, Chengdu and Guiyang) in central and western provinces are detected as HL cluster areas (pink parts). These cities act as growth poles in these provinces, with adequate development endowments compared to the neighboring cities. The green parts are LH clusters, including those areas surrounding HH clusters like the western and eastern part of Guangdong, the northern part of Juangsu and the southern part of Zhejiang. LL clusters are covered by blue parts. The results show that the northeast regions, the western regions and parts of the central regions of China are less attracted to highly educated talents.

Determinants and Spatial Effects of Highly Educated Talents

Variable Selection Strategy

The dependent variable of this study is the number of highly educated talents in each city in 2015, while the independent variables mainly include economic opportunity variables, amenity variables, and control variables (Table 1). In terms of economic opportunity variables, regional GDP (GDP) is the representative of regional economic development level and the future development prospect of a region, which many consider the main economic driver affecting the distribution of talents. In addition, average wage (WAGE) and unemployment rate (UNEMP) of a region are also important factors affecting the settlement of talents. With the development of the economy, the labor force will shift from primary and secondary industries to tertiary industry, thus industrial structure (INDUSTRY) may also exert a certain impact on talents. In summary, we selected the above five variables to represent economic opportunities.

Table 1 Description of variables and expected effect

Variables	Descriptions	Expected effect
Dependent variables:		
TALENT	Number of college graduates or above in each city in 2015	
Economic opportunities:		
GDP	Total GDP of each city in 2014 (Ten Thousand Yuan)	+
WAGE	Average wages of employees in urban areas in cities and towns in 2014 (Yuan)	+
UNEMP	Urban unemployment rate in 2014 (%)	–
INDUSTRY	Proportion of tertiary industry in GDP in each city in 2014 (%)	+
Amenities:		
SEBYEXP	Proportion of science and education expenditure in per capita fiscal expenditure of each city in 2014 (%)	+
EXPBYINC	Proportion of per capita financial expenditure in per capita fiscal revenue of each city in 2014 (%)	+
TEACHER	Number of primary school teachers per 100,000 primary school students in each city in 2014	+
GREEN	Greening rate of built-up areas in each city in 2014 (%)	+
WATER	Sewage treatment compliance rate in each city in 2014 (%)	+
SO2	Total industrial sulfur dioxide emissions in each city in 2014 (Ton)	–
DOCTOR	Number of doctors per 10,000 people in each city in 2014	+
ROAD	Road Accessibility Index for each city in 2014	+
RAIL	Railway passengers in each city in 2014 (10,000 person)	+
AMENITY	Comfort Index in each City in 2014	–
CONSUME	Number of star-rated hotels in each city in 2013	+
Control variables:		
DENSITY	Population density of each city in 2014 (Persons per Square Kilometer)	+
UNISTU	Number of college students per 10,000 in each city in 2014	+
FINVEST	Per capita fixed asset investment in cities in 2014 (10,000 Yuan)	Unknown
PROVINCE	Whether it is a provincial capital, a municipality or a separate city	+

According to Gleaser et al.'s (Glaeser et al. 2001) definition, amenities include four parts: a good natural environment, a supply of public services, a high speed of transportation and communication, and a variety of consumer facilities. With the improvement of people's living standards, talents are paying increasing attention to the comfort of the natural environment. The following formula gives the comfort index of each city:

$$I = T - 0.55(1 - RH)(T - 58) \tag{5}$$

where I denotes the comfort index, T is the temperature (Fahrenheit), and RH is the relative humidity (%). Bai et al. (2009) further divided the human comfort level into nine levels, and they considered 65 as the absolute comfort value. We constructed a continuous comfort index, calculating the distance from each city's comfort index to the absolute comfort value as a quantitative indicator of urban comfort. The smaller the distance, the more comfortable this city is.

Finally, we selected four indicators to evaluate the comfort of the natural environment, including air quality (SO2), water quality (WATER), urban greening rate (GREENING), and comfort index (AMENITY).

Furthermore, public services include basic services such as medical care and education. In China, public services and their infrastructure provide a good guarantee for talents to work and live. We selected the ratio of per capita science and education expenditure (SEBYEXP), the ratio of per capita fiscal expenditure (EXPBYINC), the number of teachers per 10,000 primary school students (TEACHER), and the number of doctors per capita (DOCTOR) to measure the impact of public services on the distribution of talents.

Glaeser et al. (2001) believed that contemporary urban growth depends on its ability to provide services and consumer goods. We used the number of star-rated hotels (CONSUME) as a proxy variable to measure the ability of a city to provide consumer services. In addition, transportation is an important medium for the mobility of labor. Highways and railways are the commonest means of transportation for talents. We used road accessibility (SPEED) and the number of railway passengers (RAIL) to reflect a city's traffic development, within which sDNA software calculated road accessibility as follows:

$$SPEED(x) = \sum_{y \in Rx} \frac{p(y)}{d(x, y)} \tag{6}$$

where $p(y)$ is the weight of node y in the radius R. In continuous space analysis, $p(y) \in [0,1]$; $d(x,y)$ is the shortest topological distance from node x to node y; $SPEED(x)$ is the accessibility index. We set 100 km as the radius.

In addition, some factors have a latent impact on the distribution of talents, including population density (DENSITY), fixed investments (FINVEST), administrative factors (PROVINCE), and talent supply (UNISTU). Thus, we included four control variables in our model.

Model Processing

First, we tested whether there was significant spatial autocorrelation in the model error terms based on a classical OLS regression model. The result shows that the residuals still have strong spatial autocorrelation, which confirms our speculation. The existence of spatial autocorrelation in the error terms violates the assumption of i.i.d. in the model. To solve this problem, it is necessary to use a spatial econometric model considering spatial autocorrelation.

We used GeodaSpace software to support the selection and estimation of spatial econometrics models. First, we added the above 19 explanatory variables to an OLS model, and we tested its spatial spillover effect and possible sources (Table 2). The results showed that there is a spatial dependence effect in the model residual, and the Lagrange Multiplier and Robust LM tests for spatial lag and spatial error effects were both significant, indicating that these two mechanisms may be leading to spatial autocorrelation. In addition, we tested for heteroscedasticity. The results of the Breucsh-Pagan test and the Koenker-Bassett test were significant, indicating that the data contained strong heteroscedasticity. Therefore, we used robust standard errors in estimating coefficients. Finally, we performed a strict multicollinearity test. The results showed that there was no strict multicollinearity

Table 2 Tests for spatial econometrics model

Tests	Coefficient	*P* value
Moran's I (error)	8.484***	0.000
Lagrange Multiplier (lag)	11.498***	0.001
Robust LM (lag)	8.521***	0.004
Lagrange Multiplier (error)	59.046***	0.000
Robust LM (error)	56.070***	0.000
Lagrange Multiplier (SARMA)	67.567***	0.000

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

problem in the data. According to the above tests, the application of GSM is suitable for our case.

The regression strategy was as follows. Model 2 only included four variables that reflect economic opportunities. Model 3 only included 11 variables that reflect amenities. Model 4 incorporated both economic opportunity variables and amenity variables. We further added control variables to Model 5 at the end of model process. Through this kind of strategy, it is possible to examine the mechanism of economic opportunity and amenity factors to some extent.

Result Analysis

The regression results are in Table 3. In Model 2, the economic opportunity variables (except for UMEMP) had a significant impact on the number of highly educated talents. Specifically, improvement in regional development level and regional wage level, and the optimization of the regional industrial structure both increased the number of highly educated talents. From the perspective of the spatial spillover effect, both the spatial lag effect of the dependent variable (ρ) and the spatial error effect of the model error term (λ) were significant, but the two influences were opposite: an increase in the number of talents in a region resulted in a decrease in talents affected by spatial lag effect, while an increase of talents affected by spatial error effects was also significant. This is because in increasingly fierce talent competition, an increase in the number of talents in a certain city will likely inhibit the number of talents in its surrounding area, resulting in a negative coefficient of spatial lag. However, other factors may lead to a positive impact of talents in one city and its neighbors, such as cross-regional talent policies, positive policy responses of neighboring cities to a certain city that has already promulgated a talent policy, or a cross-city social network of talents.

In Model 3, when we added the amenity variables, we found that public service variables, such as SEBYEXP, EXPBYINC, and TEACHER were not significant. In addition, environment variables such as GREENING, SO2, and AMENITY did not meet our expectations. An increase in the greening rate and the improvement of natural comfort significantly inhibited the number of talents, while an increase in sulfur dioxide emissions promoted the number of talents. This is because, if we do not control for the effects of economic factors, environment variables, especially those variables representing pollution, have close relationships with the level of regional economic development. Unlike public and environment variables, traffic variables such as SPEED and RAIL were

Table 3 Results of OLS model and GSM models

	Model 1 OLS	Model 2 GSM	Model 3 GSM	Model 4 GSM	Model 5 GSM
	ln(TALEN)	ln(TALENT)	ln(TALENT)	ln(TALENT)	ln(TALENT)
ρ		−0.248**	−0.254**	−0.267***	−0.106
		(0.111)	(0.137)	(0.099)	(0.092)
λ		0.731***	0.532***	0.758***	0.683***
		(0.050)	(0.087)	(0.047)	(0.054)
ln(GDP)	0.371***	0.548***		0.552***	0.525***
	(0.066)	(0.048)		(0.055)	(0.059)
ln(WAGE)	0.528**	0.811***		0.809***	0.768***
	(0.246)	(0.216)		(0.250)	(0.253)
UNEMP	−0.007	0.003		−0.008	−0.008
	(0.014)	(0.009)		(0.008)	(0.009)
INDUSTRY	0.008*	0.019***		0.016***	0.011***
	(0.004)	(0.003)		(0.003)	(0.004)
SEBYEXP	0.006		0.010	−0.014**	−0.010
	(0.008)		(0.009)	(0.007)	(0.007)
EXPBYINC	0.001**		−0.000	0.001***	0.001***
	(0.000)		(0.000)	(0.000)	(0.000)
ln(TEACHER)	0.224		−0.038	0.327**	0.405**
	(0.183)		(0.187)	(0.147)	(0.163)
GREEN	−0.007**		−0.008*	−0.002	−0.002
	(0.004)		(0.004)	(0.003)	(0.003)
WATER	−0.002		−0.002	−0.002	−0.002
	(0.003)		(0.002)	(0.002)	(0.002)
ln(SO2)	0.038		0.158**	0.039	0.043
	(0.041)		(0.054)	(0.032)	(0.034)
ln(DOCTOR)	0.305***		0.476***	0.115*	0.174**
	(0.089)		(0.107)	(0.063)	(0.071)
ln(SPEED)	0.170**		0.402***	0.248***	0.219***
	(0.067)		(0.107)	(0.061)	(0.071)
ln(RAIL)	−0.011		0.042**	−0.013	−0.014
	(0.017)		(0.215)	(0.011)	(0.012)
ln(AMENITY)	0.004		0.009*	0.003	0.003
	(0.003)		(0.005)	(0.004)	(0.003)
CONSUME	0.001		0.005**	0.000	0.000
	(0.001)		(0.002)	(0.000)	(0.001)
Control variables	Yes	No	No	No	Yes
CONSTANT	−1.354	−2.684	13.446***	−4.582	−4.756
	(2.916)	(3.442)	(2.454)	(3.881)	(3.559)
N	224	224	224	224	224
Pseudo R^2	0.722	0.616	0.568	0.658	0.707
Standard error	Robust	Robust	Robust	Robust	Robust

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

positively associated to the number of talents. Furthermore, CONSUME also had a positive impact on talents. As in Model 2, the spatial lag effect was significantly negative, while the spatial error effect was positive, and the coefficient of spatial error term was greater than that of the spatial lag term.

Model 4 considered economic opportunity variables and amenity variables together. We found that the impact of economic opportunity variables was still robustly significant. After controlling for the impacts of economic factors, environment-related variables (e.g., SO2, GREENING, WATER, and AMENITY) did not have significant impacts of the number of talents. However, among public service variables, in line with our expectations, EXPBYINC, TEACHER, and DOCTOR had links with the number of talents. Inconsistent with our expectations, SEBYEXP had a negative correlation with the number of talents. This may be because this variable only represents the degree to which the regional government attaches importance to science and education, but it does not represent the actual amount of science and education expenditure. In addition, the role of SPEED in the region was still significant. In terms of spatial overflow variables, like Model 2 and Model 3, ρ and λ were negative and positive, respectively.

Finally, Model 5 further incorporated relevant control variables, including per capita fixed asset investment, population density, the number of college students, and city administrative level. The results showed that regional gross GDP and average wage still had a significant impact on the number of talents, and their coefficients were relatively high. In addition, one cannot ignore the impact of industrial structure. In terms of amenity factors, basic public services such as education and medical care had significant associations with talents, and road accessibility in a region was positively related to the number of talents. After adding the control variables, AMENITY and CONSUME did not show significant relationships with talent distribution.

On the other hand, when we added the control variables, the spatial lag term did not pass the significance test any more, while the spatial error term was still robustly significant, and the coefficient reached 0.683. According to Model 5, it seems that the space spillover effect of highly educated talents in Chinese cities comes mainly from the impacts of the spatial error term. To be specific, it comes from influences of several unobserved factors such as cross-regional or country specific policies, policy responses, social network linkages. With a background of talent competition, cities with talent policies will more or less affect the policy-making of their surrounding cities, thus promoting the number of talents in their neighborhoods. In addition, the promotion of cross-city or country specific public service policies (e.g., Internet plus policy) has led to a reduction in barriers to the interactions between cities, leading to the spillover of the spatial distribution of talents (Miguélez et al. 2010). Last but not least, the cross-city social networks talents build also matter. The linkage of talents in neighboring cities can help to offset the risks of integration and employment for talents (Gu et al. 2018; Miguélez et al. 2010).

In general, the economic opportunity variables were significant in Model 2, Model 4, and Model 5, and the coefficients were relatively large, showing a certain degree of robustness. In terms of amenity variables, a considerable number of variables did not meet the expectations in Model 2. After considering the impacts of economic opportunities, the results show that public service variables and traffic variables significantly affected the gathering of talents, whereas there is no evidence showing that regional natural comfort and consumption made a difference. In summary, economic opportunities mainly affected the spatial cluster of talents, although one cannot ignore the impacts of public services and accessibility in amenity factors. The demand for talents of economic conditions was more fundamental, and only after the economic conditions are right will non-movable amenities such as public services become important forces for talents.

Conclusion

We first used GIS spatial analysis methods to explore the spatial distribution pattern of highly educated talents at the prefecture level. Second, we employed a GSM model to discuss the driving roles of economic opportunities and amenity in the distribution of talents, and we discussed the spatial spillover effects of talents. Our conclusions are as follows:

First, the spatial pattern of highly educated Chinese talents is highly concentrated and unbalanced. Cities with higher administrative levels and larger populations have more talents living there. The *MC* of the density and quantity of highly educated talents in prefecture-level cities in China is significantly positive, indicating that there is a cluster pattern of talents, where the spatial spillover effect is obvious. The result of the LISA analysis shows that the urban agglomerations are HH clusters of talent density, and a few inland growth poles are covered by HL clusters. Cities surrounding to HH clusters are detected as LH clusters, and the southeastern regions, western regions and part of central regions of China are less attracted to Talents.

Second, economic opportunities are still the main forces affecting the cluster patterns of highly educated talents. Economic opportunity variables have significant positive impacts on the number of talents. Among them, wages are the core driving force. The distribution and migration of talents depend on rational decisions involving economic opportunities. In addition, variables like regional development level and industrial structure matter.

Third, when controlling for the influence of economic opportunities, amenity variables (e.g., basic public services and accessibility of transportation) play significant roles in the spatial agglomeration of highly educated talents. Only when the requirements for economic opportunities are right will amenities work. When governments do not consider economic opportunities, the impacts of amenities on talents is not clear. In addition, variables such as natural comfort and consumption ability have no significant effect on talent agglomeration.

Fourth, the spatial spillover effect of China's highly educated talents mainly comes from the influences of cross-regional or country specific talent, economic development policies, and social network linkages effects of talents. For a city, a talent policy may lead to a positive response from its surrounding cities, which will lead to an increase of talents in these cities. In addition, the promotion of cross-city public services has led to a reduction in barriers for talents to communicate between cities. Cross-city social networks talents build in neighboring cities have similar effects.

Discussion

Generally speaking, the unbalanced patterns of cities' economic opportunity factors and amenity factors are the main reasons for the uneven distribution of Chinese talents. In recent years, the trend of highly educated talents gathering together is obvious. The southeastern coastal areas have attracted a large percentage of talents because of its developed economic level and the high amenities in these regions. On the one hand, the concentration of talents has had a large effect in these developed regions, and this has strengthened the knowledge spillover effect, which has stimulated the innovation and economic development of these regions (Zeng et al. 2019). On the other hand, despite the agglomeration of talents in developed areas, there is a lack of talents in the central and western regions, which has squeezed the development space of these regions.

Having realized the unbalance spatial distribution of highly educated talents, the central government should optimize this pattern by formulating appropriate talent development policies, and it should improve the efficiency of talent allocation. For highly concentrated areas of talents, we must pay attention to their positive externalities, and maybe it is not appropriate to intervene excessively through policies. However, for the central and western regions, where talent distribution is sparse, it may be necessary to formulate talent policies to give certain economic rewards to young graduates to guarantee their basic living conditions. Local governments should support preferential economic policies, such as returning hometown subsidies, returning to the countryside, and reemployment economic subsidies, and they should encourage migrants working in developed areas to return to their hometowns to work and start businesses, thereby narrowing the regions' differences in China.

Finally, we must pay more attention to the spatial spillover effect of talents in neighboring regions. Governments should formulate a series of coordinated talent policies to strengthen the cooperation and the joint driving role within the region, considering that adjacent cities and regions may have natural geographical advantages, similar cultural genes, and adjacent social networks. The application of integrated regional talent policies is conducive to stimulating the coordinated development of regions.

Our paper still has the following shortcomings. First, the empirical study used survey data, and there is a problem with missing variables. In particular, there are some key factors we cannot include in the variables of amenities that may affect the concentration of talents. Second, this paper only builds a cross-sectional model based on the 2015 census data, and it lacks panel empirical research. Third, regarding the definition of talents, maybe highly educated talents cannot fully represent talent groups. The extent to which our research results can explain the distribution of talents needs further consideration. In the future, we will carry out research on the following aspects. We should further enrich talent data sources through field research or big data platforms to improve the accuracy of our results.

Acknowledgements This study was funded by the National Natural Science Foundation of China (No.71733001), the National Social Science Foundation of China (No.17ZDA055), and the foundation of National Academy of Innovation Strategy (No.CXY-ZKQN-2019-04).

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis, 27*(2), 93–115.

Arntz, M. (2010). What attracts human capital? Understanding the skill composition of interregional job matches in Germany. *Regional Studies, 44*(4), 423–441.

Bai, Q., Huo, Z., He, N., Li, N., & Wei, N. (2009). Analysis of human body comfort index of 20 tourist cities in China. *Chinese Journal of Ecology, 28*(8), 1607–1622.

Bao, H., Fu, G., & Zhu, Y. (2007). Comprehensive evaluation of talent environment in western China – based on factor analysis. *Northern Economy, 2007*(13), 44–46.

Cebula, R. J., & Vedder, R. K. (1973). A note on migration, economic opportunity, and the quality of life. *Journal of Regional Science, 13*(2), 205–211.

Chen, Y., & Rosenthal, S. S. (2008). Local amenities and life-cycle migration: Do people move for jobs or fun? *Journal of Urban Economics, 64*(3), 519–537.

Chiquiar, D., & Hanson, G. H. (2005). International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States. *Journal of Political Economy, 113*(2), 239–281.

Clark, T. N., Lloyd, R., Wong, K. K., & Jain, P. (2002). Amenities drive urban growth. *Journal of Urban Affairs, 24*(5), 493–515.

Davies, P. S., Greenwood, M. J., & Li, H. (2001). A conditional logit approach to US state-to-state migration. *Journal of Regional Science, 41*(2), 337–360.

Deng, J., Yin, T., & Zhou, L. (2001). Research on the strategies and determinants of the development of the talents in Guangzhou. *South China Population, 2001*(4), 29–33.

Duan, D., Du, D., Liu, C., & Grimes, S. (2016). Spatio-temporal evolution of urban innovation structure based on zip code geodatabase: An empirical study from Shanghai and Beijing. *Journal of Geographical Sciences, 26*(12), 1707–1724.

Esenwein-Rothe, I. (1964). Population redistribution and economic growth, United States 1870–1950. *Jahrbücher für Nationalökonomie und Statistik, 176*(1), 150–155.

Faggian, A., Rajbhandari, I., & Dotzel, K. R. (2017). The interregional migration of human capital and its regional consequences: A review. *Regional Studies, 51*(1), 128–143.

Ferguson, M., Ali, K., Olfert, M. R., & Partridge, M. (2007). Voting with their feet: Jobs versus amenities. *Growth and Change, 38*(1), 77–110.

Florida, R. (2002). *The rise of the creative class* (Vol. 9). New York: Basic books.

Florida, R., Mellander, C., & Qian, H. (2012). China's development disconnect. *Environment and Planning A, 44*(3), 628–648.

Fujita, M., & Thisse, J. F. (1996). Economics of agglomeration. *Journal of the Japanese and international economies, 10*(4), 339–378.

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. *Journal of Economic Geography, 1*(1), 27–50.

Gottlieb, P. D., & Joseph, G. (2006). College-to-work migration of technology graduates and holders of doctorates within the United States. *Journal of Regional Science, 46*(4), 627–659.

Graves, P. E. (1976). A reexamination of migration, economic opportunity, and the quality of life. *Journal of Regional Science, 16*(1), 107–112.

Graves, P. E. (1979). A life-cycle empirical analysis of migration and climate, by race. *Journal of Urban Economics*, (*6*, 2), 0–147.

Greenwood, M. J., & Hunt, G. L. (1989). Jobs versus amenities in the analysis of metropolitan migration. *Journal of Urban Economics, 25*(1), 1–16.

Gu, C., Cai, J., Zhang, W., Ma, Q., Chan, R. C. K., Li, W., & Shen, D. (1999). A study on the patterns of migration in Chinese large and medium cities. *Acta Geographica Sinica, 1999*(3), 14–22.

Gu, H., Xiao, F., Shen, T., & Liu, Z. (2018). Spatial difference and influencing factors of settlement intention of urban floating population in China:Evidence from the 2015 national migrant population dynamic monitoring survey. *Economic Geography, 38*(11), 22–29.

Gu, H., Shen, T., Liu, Z., & Meng, X. (2019a). Driving mechanism of interprovincial population migration flows in China based on spatial filtering. *Acta Geographica Sinica, 74*(2), 222–237.

Gu, H., Liu, Z., & Shen, T. (2019b). Spatial pattern and determinants of migrant workers' interprovincial hukou transfer intention in China: Evidence from a National Migrant Population Dynamic Monitoring Survey in 2016. *Population, Space and Place*, e2250.

Knapp, T. A., & Gravest, P. E. (1989). On the role of amenities in models of migration and regional development. *Journal of Regional Science, 29*(1), 71–87.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy, 99*(3), 483–499.

Lee, S. Y., Florida, R., & Acs, Z. (2004). Creativity and entrepreneurship: A regional analysis of new firm formation. *Regional Studies, 38*(8), 879–891.

Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The Manchester School, 22*(2), 139–191.

Li, X., Fan, M., Yang, Z., & Guo, L. (2018). Influencing factors of the development environment of science and technology talents based on Struct. *Forum on Science and Technology in China, 2018*(8), 147–154.

Li, R., Wu, D., Bao, J., Qiu, Y., & Wang, W. (2013). Evolution of temporal and spatial pattern and driving mechanisms of agglomerative growth of senior science talents: A case study of academicians of CAS. *Progress in Geography, 32*(7), 1123–1138.
Liu, Z., & Gu, H. (2019). Evolution characteristics of spatial concentration patterns of interprovincial population migration in China from 1985 to 2015. *Applied Spatial Analysis and Policy*, 1–17.
Liu, Y., & Shen, J. (2014). Jobs or amenities? Location choices of interprovincial skilled migrants in China, 2000–2005. *Population, Space and Place, 20*(7), 592–605.
Liu, Y., Shen, J., Xu, W., & Wang, G. (2017). From school to university to work: Migration of highly educated youths in China. *The Annals of Regional Science, 59*(3), 651–676.
Lucas, R. E., Jr. (1988). On the mechanics of economic development. *Journal of Monetary Economics, 22*(1), 3–42.
Luo, S. (2001). The mechanisms, effects and countermeasures of Chinese university graduates' transregional flow. *Soft Science, 2001*(2), 84–88.
Miguélez, E., Moreno, R., & Suriñach, J. (2010). Inventors on the move: Tracing inventors' mobility and its spatial distribution. *Papers in Regional Science, 89*(2), 251–274.
Miu, G., Chen, W., & Tang, C. (2013). Study on influence factors of talent agglomeration in high-tech enterprises——Taking commercial Aircraft Corporation of China ltd for example. *Science and Technology Management Research, 33*(10), 120–122.
Mueser, P. R., & Graves, P. E. (1995). Examining the role of economic opportunity and amenities in explaining population redistribution. *Journal of Urban Economics, 37*(2), 176–200.
Niedomysl, T., & Hansen, H. K. (2010). What matters more for the decision to move: Jobs versus amenities. *Environment and Planning A, 42*(7), 1636–1649.
Rodríguez-Pose, A., & Tselios, V. (2011). Mapping the European regional educational distribution. *European Urban and Regional Studies, 18*(4), 358–374.
Romer, P. M. (1990). Endogenous technological change. *Journal of political Economy, 98*(5, Part 2), S71–S102.
Scott, A. J. (2010). Jobs or amenities? Destination choices of migrant engineers in the USA. *Papers in Regional Science, 89*(1), 43–63.
Song, H., Zhang, M., & Wang, R. (2016). Amenities and spatial talent distribution: Evidence from the Chinese IT industry. *Cambridge Journal of Regions, Economy and Society, 9*(3), 517–533.
Stark, O., & Bloom, D. E. (1985). The new economics of labor migration. *The American Economic Review, 75*(2), 173–178.
Stillwell, J., Daras, K., & Bell, M. (2018). Spatial aggregation methods for investigating the MAUP effects in migration analysis. *Applied Spatial Analysis and Policy, 11*(4), 693–711.
Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography, 46*(sup1), 234–240.
Todaro, M. P. (1969). A model of labor migration and urban unemployment in less developed countries. *The American Economic Review, 59*(1), 138–148.
Wang, W., & Zhang, S. (2003). Ponder on character and reason of our country talents mobility at present. *Journal of Chongqing Vocational& Technical Institute, 2003*(1), 18–20.
Xu, Z., & Ouyang, A. (2018). The factors influencing China's population distribution and spatial heterogeneity: A prefectural-level analysis using geographically weighted regression. *Applied Spatial Analysis and Policy, 11*(3), 465–480.
Xu, Q., & Zhang, T. (2010). Talent agglomeration based on City environment. *China Population, Resources and Environment, 20*(9), 171–174.
Yu, Z., Zhang, H., Tao, Z., & Liang, J. (2019). Amenities, economic opportunities and patterns of migration at the city level in China. *Asian and Pacific Migration Journal*, 0117196819832054.
Yuan, X., & Wan, N. (2007). Floating population, informal employment and development of large cities. *Journal of China University of Geosciences (Social Sciences Edition), 2007*(5), 28–33.
Zeng, J., Liu, Y., Wang, R., & Zhan, P. (2019). Absorptive capacity and regional innovation in China: An analysis of patent applications, 2000–2015. *Applied Spatial Analysis and Policy*, 1–19.
Zhang, J., Du, D., & Jiang, H. (2011). Regional differences of S&T Talents in Jiangsu Province. *Scientia Geographica Sinica, 31*(3), 378–384.
Zhang, W., Jing, W., & Wang, Y. (2017). What determines the talent attractiveness of first-tier cities? Based on the random forest algorithm of influential factors. *Science and Technology Management Research, 37*(22), 99–108.