Correlation between Population Structure and Regional Innovation Ability Based on Big Data Analysis

Lulu Qin, Ruofei Lin, Rongwei Gao, and Han Lin

1 College of Design and Innovation, Tongji University, Shanghai 200092, China
2 School of Economic and Management, Tongji University, Shanghai 200092, China
3 Shanghai International College of Intellectual Property, Tongji University, Shanghai 200241, China
4 Department of Urban Planning and Design, Xi’an Jiaotong-Liverpool University, Suzhou, Jiangsu 215000, China

Correspondence should be addressed to Rongwei Gao; rongwei.gao@tongji.edu.cn and Han Lin; han.lin19@student.xjtlu.edu.cn

Received 21 June 2022; Revised 13 August 2022; Accepted 17 August 2022; Published 31 August 2022

Academic Editor: Ning Cao

Copyright © 2022 Lulu Qin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the continuous advancement of the urbanization process, the scale of cities has expanded rapidly, and the amount of floating population has grown rapidly. Big data not only brings changes in thinking, technology, and management but also promotes the renewal and development of social governance concepts, technologies, methods, and models, which creates new opportunities for urban floating population governance innovation. City governments at all levels are facing enormous management pressure and urgently need to promote innovative reforms in the governance of urban floating population. The development of information technology has developed into the era of big data, and strong data collection and processing capabilities can greatly improve the scientific level of urban floating population management. In the research of spatial correlation of innovation ability, the research of domestic and foreign scholars has been very mature, and the methods used are more comprehensive, combining big data with urban floating population governance, placing urban floating population governance in the context of big data research and analysis, exploring the channels for the government to use big data to optimize urban floating population governance, and putting forward specific and feasible countermeasures and suggestions. After a series of tests, the double-fixed-effects spatial Dubin model is used to explore the influence of population structure on regional innovation ability from five aspects: population urban-rural structure, population industrial structure, population education structure, population age structure, and population density. The results show that the population structure of higher education has the most significant role in promoting regional innovation capability, followed by urban population, population density, and secondary industry population structure; the working-age population structure has no significant impact on regional innovation capability; provincial innovation capability has a significant effect positive spatial spillover effect.

1. Introduction

City is a coupling body of regional information flow, material flow, energy flow, and knowledge flow. It is a community gathering place that is different from natural settlements in rural areas. With the acceleration of globalization and urbanization, the rapid expansion of urban space, and the coexistence of internal reorganization, the stratification of urban population will also bring about a new spatial order [1]. Cities of different scales and grades have different evolution characteristics of urban space. Megacities have often experienced suburbanization, while small and medium-sized cities still rely on centripetal force as the main driving force for urban development. Under the background of high-speed, complex and diverse urban space evolution, mining spatial big data with geographic attributes and in-depth and rapid exploration of the distribution and utilization of urban living space has become an urgent need to optimize urban living space [2]. The spatial organization of residence and employment is an important element of urban spatial structure. With the expansion of urban space and the development of industries and residences in the suburbs, the problem of separation of work and housing is becoming more and more serious, which leads to a relative increase in...
commuting time and commuting distance, causing traffic congestion, environmental pollution, and increased travel costs to a certain extent [3]. The increase of urban population means the increase of traffic volume, and the speed of road construction and public transportation investment lags behind the speed of urbanization to a certain extent [4]. The separation of work and residence leads to serious tidal traffic phenomenon, causing serious traffic congestion, causing continuous deterioration of urban life and ecological environment, and becoming an “urban stubborn disease” that hinders the healthy development of cities [5].

Studying the layout of urban residential space, matching urban industrial space with residential space, and rationally arranging traffic are essential for building an eco-city [6]. In the era of rapid development of information technology, the lifestyles and life concepts of urban residents have undergone tremendous changes, and the urban spatial organization has also changed [7]; traditional urban spatial behavior research can no longer cover the complexity of today’s urban residents’ spatial behavior, and to a certain extent, the accuracy of its research and statistical results is reduced by the influence of visitors. The research data of residential space research often come from traditional data such as government statistics, planning data, and land remote sensing data. To analyze urban spatial structure through traditional urban research and survey methods, data always lag behind urban development to a certain extent and cannot capture the latest development trends in the context of rapid urbanization. Moreover, the spatial scale of the study is large, and the time span is long. The cities studied are mostly large cities, and there is a lack of research on medium and small cities. In contrast, the carried big data of geospatial attributes provides rich, detailed, and real-time data, which undoubtedly provides new opportunities for urban research and planning that focus on space [8]. Urban living space is a crucial part of urban social space [9], and exploring the distribution characteristics and utilization of urban living space is an important theoretical and practical issue faced by urban research and construction. The development of contemporary cities is changing rapidly, and timely and accurate understanding of urban spatial behavior of urban residents is crucial to the healthy development of cities. The emergence of big data technology has brought new opportunities and challenges to urban spatial research and urban planning. Mianyang City is a regional central city in the Chengdu-Chongqing urban agglomeration Chengde Mianle City [10], and the rational organization of urban space has become the guarantee of its functions in the Chengdu-Chongqing urban agglomeration.

In the research of spatial correlation of innovation ability, the research of domestic and foreign scholars has been very mature, and the methods used are more comprehensive. There are still obvious deficiencies in the research on the impact of population structure on regional innovation capacity. Most studies focus on the effects of a single factor such as the aging population, educated population, and population density on technological innovation and technological innovation, and the research perspective is not comprehensive. When studying the impact of population structure on regional innovation capabilities, most of the traditional measurement methods are used, only considering the linear structure without considering the spatial factors, which will cause endogeneity problems, which will easily lead to deviations in the model estimation results and ignore the regional innovation activities in each region. Therefore, this paper uses the super-efficiency data envelopment analysis model to measure regional innovation capabilities and fully considers the differences and dynamics of development between different regions. On the basis of testing the spatial autocorrelation of regional innovation capabilities, a spatial measurement model established to this paper empirically analyzes the impact of various demographic factors on China’s regional innovation capabilities and provides decision-making references for promoting the improvement of innovation levels from the perspective of population and collaborative innovation. This paper constructs an evaluation index system of innovation capability and measures regional innovation capability through data envelopment analysis; secondly, it analyzes the changing trend and spatial distribution characteristics of innovation capability measurement value; finally, it tests regional innovation capability with the help of Moran and Girley index.

2. Related Work

There are many ways of categorizing population structure. Usually, according to the different characteristics of the population, it can be divided into the geographical structure of the population, the social structure, and the natural structure [11]. This paper selects independent variables from the four aspects of the population’s geographical, economic, social, and natural structure to empirically analyze the changing trends and distribution characteristics of China’s population structure and its impact on regional innovation capabilities. The geographical structure of the population refers to the distribution of the population in different geographical areas [12]. It mainly includes urban and rural structure and population density. Among them, the urban-rural structure of the population is divided into the proportion of urban population and the proportion of rural population according to the living area of the population, which is an important indicator reflecting the level of urbanization [13]. Population density is the number of people living on a unit area of land, indicating the concentration of the population. It is usually greatly affected by the level of regional economic development and the environment. These effects cause differences in population density in different regions. Population economic structure refers to the division according to certain economic indicators, which includes population industrial structure and occupational structure [14]. The industrial structure of population is defined as the distribution and proportion of human resources in the primary, secondary, and tertiary industries in a certain region. Population occupational structure refers to the distribution and proportion of human resources in different occupations [15].

Population and social structure refers to dividing the population into educational structure, ethnic structure, marital status structure, etc., and their respective proportions according to the characteristics of population and
society [16]. Among them, the educational structure refers to the proportion of the population with different educational levels in the total population in a certain region; the ethnic structure refers to the proportion of the population of each ethnic group in the total population [17]. The natural structure of the population is divided into gender structure and age structure according to the physiological characteristics of the population and the respective proportions of each category. Among them, the gender structure refers to the proportion of male and female population in the total population. The age structure refers to the proportion of the population of different age groups in the total population in a certain region, which is greatly affected by migration and the natural growth rate of the population. The age structure of the population determines the future population development trend and speed of a country or region [18]. To sum up, the content involved in the population structure is complex. Based on the empirical evidence of the impact of the population structure on the regional innovation ability, this paper formulates a series of population policies to improve the regional innovation ability. Select the principles of the significance of the impact of regional innovation capabilities [19].

According to the empirical needs, this paper selects the urban-rural structure, industrial structure, education structure, age structure, and population density of the population from the geographical structure, economic structure, social structure, and natural structure of the population for research. American demographer Frank Nottestein revealed the three-stage population transition theory [20]. The demographic characteristics of the first stage are as follows: the population presents a high birth rate and death rate. Since the rate of change of the former is smaller than that of the latter, the natural increase in the mortality rate has a greater impact. Since the latter has decreased more, the natural growth rate has increased, and the age structure of the population has changed; the third stage of the population is characterized by a low birth rate and mortality rate. With less natural growth, coupled with the aging population born in the second stage, the aging rate has increased significantly. At present, the population has generally experienced a transition from high to low in birth rate, death rate, and natural growth rate. We should comprehensively analyze the reduction of the natural population growth rate and the improvement of the structure and balance the population structure. Therefore, changes in demographic structure and distribution will have an impact on regional innovation.

3. Spatial Econometric Analysis of Population Structure

3.1. Spatial Effects. The characteristic analysis of spatial effects is the premise of spatial econometrics research, and by analyzing the randomness and structure of variables, it can reveal the spatial correlation and spatial heterogeneity of variables. Spatial correlation refers to the spatial external effects generated by the proximity of observation units in geographic space and the level of socio-economic development so that observation units show spatial autocorrelation, such as the coordinated economic development of various regions in Hebei Province. Spatial heterogeneity means that, in different directions, spatial effects have spatial correlations within a certain distance range, and their influence ranges are different, or their intensity varies in magnitude. For example, the economic development of the Beijing-Tianjin-Hebei region is not synchronized. The method of spatial econometric analysis is as follows: if it is found that the explained variables have spatial correlation, a spatial econometric model can be constructed, estimated, and tested; if there is no spatial correlation, an ordinary panel model can be constructed, as shown in Figure 1. On the basis of testing the spatial autocorrelation of regional innovation capabilities, a spatial measurement model is established. This paper empirically analyzes the impact of various demographic factors.

In the spatial econometric model, it is necessary to import the spatial weight matrix used to represent the interaction between the units. The spatial weight matrix is an important tool to describe the relative position relationship of the spatial observation units and measure the spatial dependence. The commonly used spatial weight matrices are as follows:

(1) The spatial adjacency weight matrix, which assigns 0 or 1 to the elements $w_{ij}$ in the matrix according to the adjacent features between the observation units, and its $w_{ij}$ is

$$w_{ij} = \begin{cases} 
1, & i, j \in B, \\
0, & i, j \notin B.
\end{cases}$$

(2) The spatial distance weight matrix, which is weighted according to the change of the distance between observation points, whose $w_{ij}$ is

$$w_{ij} = \begin{cases} 
\frac{1}{d_{ij}}, & i, j \in B, \\
0, & i, j \notin B.
\end{cases}$$

The test of spatial autocorrelation is divided into two types: global test and local test. The global test is a measure that is used for all regions and has the same pattern or spatial process, and its measure of spatial autocorrelation is the average value of all regions. The local test is a measure of the spatial autocorrelation of each observation unit, and the measure of the observation unit at different locations is different. Moran index (Moran I) is often used in spatial correlation analysis, and the index is divided into global and local Moran I index. The global Moran I index can express the overall spatial distribution of a phenomenon in a regional unit, so as to judge whether the phenomenon has spatial aggregation, and reveal whether the attribute values in the geographical distribution have obvious spatial correlation. Its definition is as follows:
In the formula, $y_i$ is the observation value of the $i$th region, $w_{ij}$ is the spatial weight of regions $i$ and $j$, and $n$ is the number of regions. Moran $I$ takes value between $-1$ and 1. When the index value is positive, it means that there is a positive spatial correlation; when it is negative, it means that there is a negative spatial correlation; and when it is zero, it means that the space is randomly distributed. The Geary $C$ index can also test the global spatial autocorrelation. The index is obtained by using the dispersion between sample observations. The definition is shown in the following formula:

$$I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(y_i - \bar{y})^2}{s^2 \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}}. \quad (3)$$

In the formula, $y_i$ is the observation value of the $i$th region, $w_{ij}$ is the spatial weight of regions $i$ and $j$, and $n$ is the number of regions.

3.2. Measurement of Regional Innovation Capability. Therefore, in order to measure innovation ability more objectively and accurately, it is necessary to follow the principles of systematicness, scientificity, comparability, and operability in the process of constructing the evaluation index system. The operation of the innovation system is the process of inputting innovation resources and finally producing results and bringing economic benefits. Innovation input refers to innovation resources such as manpower and capital, and its size will have an impact on innovation scale, innovation output and speed, and innovation output generally includes scientific and technological achievements, economic benefits, and other aspects. If a region has less input of innovation resources but relatively more output, then the region’s innovation capability is high; if the region has more input of innovation resources and less output, it means that its innovation level is low. Referring to the previous literature on measuring regional innovation capability from the perspective of input and output, this paper finally selects relevant evaluation indicators from the following four aspects. The indicator system is shown in Figure 2.

There are 6 input indicators in the innovation evaluation system. Human resources input includes full-time equivalent of R&D personnel in industrial enterprises above designated size, reflecting the investment scale and potential of independent innovation talents in various regions; employees in scientific research, technical services, and geological surveying. The number reflects the investment of scientific research talents in various regions; the number of full-time teachers in ordinary institutions of higher learning reflects the human input of higher education in various regions. Innovation capital investment includes R&D expenditure of industrial enterprises above designated size, reflecting the scale of innovation activities in each region; expenditure of industrial enterprises above designated size for developing new products, reflecting the support of local governments for developing new products; education expenditure, reflecting each region scale of educational
activities. In the formula, $y_i$ is the observation value of the $i$th region, $W_{ij}$ is the spatial weight of regions $i$ and $j$, and $n$ is the number of regions. MoranI takes value between $-1$ and $1$.

There are 4 output indicators in the innovation evaluation system. The output of scientific and technological achievements includes the number of scientific and technological papers included in major foreign search tools, reflecting the theoretical scientific and technological achievements and research and development capabilities of innovation in various regions; the number of domestic three patent applications and authorization, reflecting the regional product output and efficiency of R&D activities. The output of economic benefits includes the technical market turnover, which reflects the exchange of innovation achievements between different regions; the sales revenue of new products of industrial enterprises above designated size, which reflects the technological innovation achievements of each enterprise and its ability to convert into economic benefits after it is introduced to the market. This chapter adopts the super-efficiency DEA model, which requires the number of decision-making units to be greater than the product of the number of input and output indicators and more than three times the sum of the number of input and output indicators. This paper selects the innovation data of 30 provinces from 2017 to 2021 as different decision-making units for efficiency measurement. The operation of the regional innovation system is a relatively complex process, and the innovation ability is affected by many factors. To measure it, it is necessary to comprehensively consider the selection of indicators to build a scientific multilevel evaluation index system.

Import the innovation input and output data into the model, use the MyDEA software to complete the calculation process, and finally get a total of 10 years of innovation ability measurement value. The measurement results are shown in Figure 3.

The national regional innovation capability showed a wave-like upward trend. From 2008 to 2009, the average regional innovation capability increased from 0.582 to 0.584, and it declined in 2020. The average regional innovation capability decreased from the previous year (0.005). Since 2011, the regional innovation capability has continued to rise. The average regional innovation capability in 2021 is 0.720, an increase of 0.138 compared with 2011. In terms of subregions, the innovation level in the eastern region has shown a wave-like upward trend. Among them, the measurement value of innovation capability in the eastern region decreased by 0.006 in 2011 and improved in other years. The central region only decreased by 0.014 in 2011, and the regional innovation capability in the western region decreased by 0.060 from 2010 to 2011. There are certain differences in the regional innovation capabilities of the eastern, central, and western regions. In 2008, the average innovation capabilities of the eastern region were 0.201 and 0.216 higher than those of the central and western regions, respectively. The innovation capability measure DEA is valid in 1 region. It shows that, with the passage of time, the gap between eastern and central and
central and western innovation capacity is gradually decreasing.

3.3. Spatial Distribution Characteristics of Regional Innovation Capabilities. According to the calculation results of regional innovation ability, use ArcGIS software to draw its spatial distribution map and use different colors to divide these provinces and municipalities into four grades according to the size of the value, which more intuitively reflects the spatial distribution characteristics of regional innovation ability and the characteristics of each province. Among them, Tibet, Hong Kong, Macau, and Taiwan are not included in the research scope, and they are displayed in white in the quantile chart. The darker the color in the figure, the higher the measurement value of regional innovation ability, and the lighter the color, the lower the measurement value of regional innovation ability. Figure 4 shows the spatial distribution of innovation capabilities in various regions in 2020 and 2021.

The provinces with relatively low regional innovation ability values are mainly distributed in the middle reaches of the Yangtze River and the greater northwest region, and the provinces with relatively high regional innovation ability values are mainly distributed in the coastal areas. It shows that the development level of regional innovation capacity in each province is extremely uneven, and adjacent provinces have the characteristics of strip and block spatial agglomeration. From a specific analysis, there are 9 provinces in the first echelon with the lowest regional innovation capability value in 2008, namely, Xinjiang, Qinghai, Inner Mongolia, Ningxia, Shanxi, Hebei, Henan, Guangxi, and Jiangxi. There are 8 provinces in the second echelon with low regional innovation capability value, namely, Liaoning, Fujian, and Hainan in the east; Heilongjiang and Jilin in the middle; and Guizhou, Yunnan, and Gansu in the west. Seven provinces belong to the third echelon with higher regional innovation capability values: Shandong on the northern coast, Shaanxi in the middle reaches of the Yellow River, Anhui, Hubei, and Hunan in the middle reaches of the Yangtze River, and Chongqing and Sichuan in the southwest. The provinces and cities that belong to the fourth echelon with the highest regional innovation capability value are all located in the coastal areas, namely, Zhejiang, Beijing, Tianjin, Jiangsu, Guangdong, and Shanghai.

The innovation capabilities of Xinjiang, Henan, Guangxi, and Jiangxi have also improved, from the first echelon with the lowest innovation capability measurement value to the second echelon with a low regional innovation capability value. Qinghai, Inner Mongolia, Ningxia, Shanxi, and Hebei have always been in the first echelon. Although the innovation capability value of these regions has increased, their growth rate is not large enough, as shown in Figure 5. The regional innovation capability of Shaanxi has dropped from the third echelon to the second echelon. Analysis of the target value of the input indicators in the DEA measurement results shows that, in 2017, Shaanxi had a phenomenon of low utilization rate of R&D personnel in terms of innovation investment, resulting in a serious waste of human resources, the output cannot achieve the expected effect, resulting in a decline in innovation capacity. Generally speaking, the regions with the highest innovation capability values are mainly concentrated in the Bohai Rim and Yangtze River Delta regions. The economic development level of these regions is relatively high, indicating that there is a certain scale of spatial agglomeration effect centered on developed provinces.

4. Global Spatial Correlation Test

Based on the spatial adjacency weight matrix of the Rook weighting method, this paper uses Stata software to calculate the global Moran I index and Geary C index to verify whether there is a spatial correlation in the regional innovation capabilities of provinces from 2017 to 2021. The test results are shown in Figure 6.

According to the test results of 30 provinces from 2017 to 2021, the Moran I index value of the global regional innovation capacity is between 0 and 1, and the gear c index is less than 1, indicating that the overall innovation capacity presents positive spatial autocorrelation characteristics. This positive spatial correlation indicates that high-innovation regions tend to be adjacent to other high-innovation regions, while low-innovation regions tend to be adjacent to other low-innovation regions, which verifies the spatial distribution of regional innovation capabilities in the quartile map. The situation where provinces with similar colors are concentrated with each other is shown in Figure 7. This is because regions with comparable innovation capabilities have certain similarities in terms of human resources, capital flows, and policies, so there are frequent interactions between these regions. From the perspective of the change trend in time, the Moran I index value shows a fluctuating upward trend as a whole, and the Geary C index shows a fluctuating downward trend. Both indexes passed the 10% significance level test, indicating that the regional innovation ability has a significant impact on the adjacent regions which have significant codirectional effects, and the trend of regional innovation capacity agglomeration is gradually strengthening over time.

According to empirical needs, this paper selects the urban population structure, secondary industry population structure, higher education population structure, working-age population structure, and population density in the population structure. The impact of capability is reflected in both innovation input and innovation environment. Cities and towns have a large stock of human capital, which can meet the needs of various professional and technical talents for innovation activities; towns can provide better information exchange platforms and infrastructure, which is conducive to the dissemination of new knowledge, new technologies, and promotion of new products. The influence of the secondary industry population structure on innovation ability reflects the stage and degree of industrial development. The secondary industry has achieved rapid development as a whole and has gradually transformed from a labor-intensive model to a technology-intensive model. Under the market competition and favorable policy
environment, industrial enterprises are also constantly innovating to obtain monopoly competitive advantages. Therefore, the population of the secondary industry is employed. Quantity is closely related to innovation capacity. The influence coefficient of urban population structure on regional innovation ability is 0.618738, and it has passed the 10% significance level test. On the one hand, the stock of human capital is the precondition for the normal operation of the innovation system and the main factor for a country or region to carry out independent innovation. On the other hand, technical imitation of adjacent areas is one of the ways to improve innovation capabilities, and the level of absorption and transformation of new technologies is closely related to the stock of human capital. As can be seen in Figure 8, in the LM test results, the four models rejected the null hypothesis of “no spatial effect of dependent variable” and “no spatial effect of residuals.” In the test results of robust LM, the common panel model and the spatial fixed effect model only rejected the null hypothesis that “the dependent variable has no spatial effect,” and neither the time fixed effect model nor the space and time fixed effect models rejected the “dependent variable has no spatial effect,” and the null hypothesis that the residual term has no spatial effect. Comprehensive consideration: the model should contain the spatial lag term, but it cannot be determined whether it contains the residual term of spatial autocorrelation, and further testing is required.

To judge whether the spatial Dobin model can be simplified, Wald and LR tests need to be performed on the two null hypotheses of the spatial Dobin model. The former is used to test whether the spatial Dobin model can be simplified to a spatial lag model, and the latter is used to test the spatial Dobin model. It can be simplified to a spatial error model, and if both hypotheses are rejected at the same time, the spatial Doberman model is used. The following three fixed-effect spatial Doberman models were constructed, and Wald and LR tests were performed on them, respectively. The results are shown in Figure 9. It can be seen from the results that, for the two hypotheses, the spatial fixed effect and the spatial Durbin model of the temporal fixed effect...
passed the Wald and LR tests with a significance level of 1%, and the spatial Durbin model of the spatial and temporal fixed effects was 10. The significance level of % passed the Wald and LR tests, so both null hypotheses were rejected and the spatial Doberman model should be used. After comprehensive analysis of all test results, the spatial Durbin model with temporal and spatial fixed effects was selected for empirical research.

It shows that every 1% increase in the urbanization rate will increase the regional innovation ability by 0.62 percentage points, and the improvement of the urbanization level can significantly promote the improvement of innovation ability. In the process of urbanization, the level of infrastructure construction is constantly improving, attracting a large number of people to migrate to urban spaces to work and live, which can promote the emergence of new industries and the dissemination of new technologies, thereby driving innovation and development. The influence coefficient of the spatial lag term of the urbanization rate was 0.349228, which did not pass the significance level test. The influence coefficient of population density was 0.380825, and it passed the test at the 10% significance level. It shows that population density has a significant positive impact on regional innovation capacity. The increase in population density can reduce the cost of information transmission, facilitate mutual communication and cooperation among various innovation entities, and thus breed new innovation achievements; at the same time, population aggregation can enable various enterprises and research institutions to obtain sufficient supply of human resources, which is conducive to improving the level of innovation. The influence coefficient of the spatial lag term of population density is −1.995027, and it passed the 5% significance level test. It shows that the population density of a region has a significant negative effect on the innovation ability of neighboring regions. The level of population density can roughly reflect the economic development status of a region. The population basically flows from underdeveloped areas to developed areas. The increase of population density in this region is not conducive

![Figure 7: Moran scatter chart of regional innovation capability.](image1)

![Figure 8: LM and robust LM test results.](image2)

![Figure 9: Wald and LR test results for the spatial Doberman model.](image3)
to the improvement of the innovation capability of adjacent regions.

5. Conclusion
The development of information technology has developed into the era of big data, and strong data collection and processing capabilities can greatly improve the scientific level of urban floating population management. Big data not only brings changes in thinking, technology, and management but also promotes the renewal and development of social governance concepts, technologies, methods, and models, which creates new opportunities for urban floating population governance innovation. The influence coefficient of the proximity effect of regional innovation capability on regional innovation capability is 0.220975, and it has passed the 1% significance level test, that is, every 1% increase in the innovation capability of the region will promote the innovation capability of its adjacent regions to increase by 0.22%. It shows that there is a positive spatial correlation between the innovation capabilities of various provinces and municipalities, and the improvement of innovation capabilities in one region can effectively promote the development of innovation capabilities in adjacent regions. This is due to the fact that innovation activities in various regions are not independent of each other, and innovation resources and technologies flow between adjacent regions, resulting in a significant positive spillover effect on regional innovation capabilities on the whole. The influence coefficient of the working-age population structure and the influence coefficient of the spatial lag term were 0.194640 and 0.051173, respectively, which did not pass the significance level test. It shows that an increase in the working-age population of a region also does not have a significant impact on the innovation capacity of the region and its neighbors. This result may be related to the fact that the original data are sample survey data.

According to the calculation results of regional innovation ability, use ArcGIS software to draw its spatial distribution map and use different colors to divide these provinces and municipalities into four grades according to the size of the value, which more intuitively reflects the spatial distribution characteristics of regional innovation ability and the characteristics of each province.

Data Availability
The data used to support the findings of this study are included within the article.

Conflicts of Interest
The authors declare that they have no conflicts of interest.

Acknowledgments
This work was supported by College of Design and Innovation, Tongji University.

References
[1] A. Acharya, S. K. Singh, V. Pereira, and P. Singh, "Big data, knowledge co-creation and decision making in fashion industry," International Journal of Information Management, vol. 42, pp. 90–101, 2018.
[2] B. T. Asheim, "Smart specialisation, innovation policy and regional innovation systems: what about new path development in less innovative regions?" Innovation: The European Journal of Social Science Research, vol. 32, no. 1, pp. 8–25, 2019.
[3] I. Castelo-Branco, F. Cruz-Jesus, and T. Oliveira, "Assessing industry 4.0 readiness in manufacturing: evidence for the European union," Computers in Industry, vol. 107, pp. 22–32, 2019.
[4] A. Ciffolilli and A. Muscio, "Industry 4.0: national and regional comparative advantages in key enabling technologies," European Planning Studies, vol. 26, no. 12, pp. 2323–2343, 2018.
[5] A. N. El-Kassar and S. K. Singh, "Green innovation and organizational performance: the influence of big data and the moderating role of management commitment and HR practices," Technological Forecasting and Social Change, vol. 144, pp. 483–498, 2019.
[6] M. Gerlach, B. Farb, W. Revelle, and L. A. Nunes Amaral, "A robust data-driven approach identifies four personality types across four large data sets," Nature Human Behaviour, vol. 2, no. 10, pp. 735–742, 2018.
[7] S. Gupta, H. Chen, B. T. Hazen, S. Kaur, and E. D. Santibañez Gonzalez, "Circular economy and big data analytics: a stakeholder perspective," Technological Forecasting and Social Change, vol. 144, pp. 466–474, 2019.
[8] Z. Lv, X. Li, H. Lv, and W. Xiu, "BIM big data storage in WebV/RGIS," IEEE Transactions on Industrial Informatics, vol. 16, no. 4, pp. 2566–2573, 2020.
[9] R. K. Mavi, R. F. Saen, and M. Goh, "Joint analysis of eco-efficiency and eco-innovation with common weights in two-stage network DEA: a big data approach," Technological Forecasting and Social Change, vol. 144, pp. 553–562, 2019.
[10] M. Mohammadpoor and F. Torabi, "Big Data analytics in oil and gas industry: an emerging trend," Petroleum, vol. 6, no. 4, pp. 321–328, 2020.
[11] K. Y. Ngiam and I. W. Khor, "Big data and machine learning algorithms for health-care delivery," The Lancet Oncology, vol. 20, no. 5, pp. e262–e273, 2019.
[12] A. M. S. Osman, "A novel big data analytics framework for smart cities," Future Generation Computer Systems, vol. 91, pp. 620–633, 2019.
[13] V. Palanisamy and R. Thirunavukarasu, "Implications of big data analytics in developing healthcare frameworks – a review," Journal of King Saud University-Computer and Information Sciences, vol. 31, no. 4, pp. 415–425, 2019.
[14] X. Pan, B. Ai, C. Li, X. Pan, and Y. Yan, "Dynamic relationship among environmental regulation, technological innovation and energy efficiency based on large scale provincial panel data in China," Technological Forecasting and Social Change, vol. 144, pp. 428–435, 2019.
[15] R. D. Raut, S. K. Mangla, V. S. Narwane, B. B. Gardas, P. Priyadarshinee, and B. E. Narkhede, "Linking big data analytics and operational sustainability practices for sustainable business management," Journal of Cleaner Production, vol. 224, pp. 10–24, 2019.
[16] S. Senthilkumar, B. K. Rai, A. A. Meshram, A. Gunasekaran, and S. Chandrakumarmangalam, "Big data in healthcare...
management: a review of literature,” *American Journal of Theoretical and Applied Business*, vol. 4, no. 2, pp. 57–69, 2018.

[17] S. A. Shah, D. Z. Seker, M. M. Rathore, S. Hameed, S. Ben Yahia, and D. Draheim, “Towards disaster resilient smart cities: can internet of things and big data analytics be the game changers?” *IEEE Access*, vol. 7, pp. 91885–91903, 2019.

[18] S. Shilo, H. Rossman, and E. Segal, “Axes of a revolution: challenges and promises of big data in healthcare,” *Nature Medicine*, vol. 26, no. 1, pp. 29–38, 2020.

[19] S. K. Singh and A.-N. El-Kassar, “Role of big data analytics in developing sustainable capabilities,” *Journal of Cleaner Production*, vol. 213, pp. 1264–1273, 2019.

[20] S. Verma, S. S. Bhattacharyya, and S. Kumar, “An extension of the technology acceptance model in the big data analytics system implementation environment,” *Information Processing & Management*, vol. 54, no. 5, pp. 791–806, 2018.