博 士 論 文

Interactive Relationship Between Urban Development and Population Mobility in China from Multi-scale Perspective

マルチスケールの視点から

中国における都市開発と人口移動の関係に関する研究

北九州市立大学国際環境工学研究科

2021 年 02 月

杨 禛

YANG ZHEN

Doctoral Thesis

Interactive Relationship Between Urban Development and Population Mobility in China from Multi-scale Perspective

February 2021

ZHEN YANG

The University of Kitakyushu
Faculty of Environmental Engineering
Department of Architecture
Gao Laboratory

Preface

This study takes Chinese cities as the research objects and explores the interactive relationship between urban development and population mobility. In particular, we emphasize the importance of studying the interaction between population mobility and urban development at different scales. Combined with official statistics and emerging big data, we describe the temporal and spatial dynamics of population mobility between cities and within cities. With the help of spatial econometric model and other mathematical statistical models, we analyze the urban development factors affecting population mobility, including urban economy, urban policies, urban built up environment and urban public services. Our empirical results provide a certain reference for the promotion of China’s urbanization process and the sustainable development of cities in the new period.

ACKNOWLEDGEMENTS

It is a great experience for me to come to Japan and follow Prof. Weijun GAO in the University of Kitakyushu. The past three years is very memorable time for me. I would like to extend heartfelt gratitude and appreciation to the people who helped me bring my study into reality.

Firstly, I would like to express the deepest appreciation to my advisor, Prof. Weijun GAO, for his great support and motivation for the work. His optimistic attitude and guidance have helped me greatly these years. In addition, Prof. Gao provides me a free space to develop my knowledge and teaches me how to become a matured researcher. Apart from the study, he also gives me much help and instruction for my daily life. I will always remember Prof. Gao's kindness.

Also, I would like to express my thanks to all university colleagues, especially Dr. Jinming JIANG, Dr. Xue FANG, Mr. Jian'an LIU, Mr. Xudong XIE, who give me the help, cooperation, and supports of research and daily life in Japan. And thanks to all the members of Gao Laboratory, especially Dr. Rui WANG, Dr. Daoyuan WEN, Mr. Fanyue QIAN, et al for their help and supporting during the study and life.

Finally, many thanks to my family for their love and encouragement which is beyond words.

Interactive Relationship Between Urban Development and Population Mobility in China from Multi-scale Perspective

ABSTRACT

Development is the primary issue facing cities in the world today. Urban development is inseparable from the support of labor. The population mobility between regions provides a guarantee for the sustainable development of the city. Therefore, the interactive relationship between urban development and population mobility requires more in-depth research. The background of the development of China's urbanization, the intersection of interdisciplinary methods and the emergence of new data enable more extensive research on this topic.

This research combines official statistics data and emerging big data to study the interactive relationship between urban development and population mobility at the macro, meso and micro levels. In addition, with the help of exploratory spatial data analysis methods, the spatial effects between urban development and population mobility, including spatial dependence and spatial heterogeneity are captured. The use of spatial econometric models reveals the driving force that affecting population mobility. The results of the empirical analysis can provide a theoretical reference for the future development of urbanization in China.

In chapter one, the background and significance of studying the interactive relationship between population mobility and urban development is demonstrated from three aspects: the development status of China's urbanization, the changes in population mobility, and the emergence of new data and new methods. Then we reviewed the previous studies. Finally, the purpose of this research is proposed.

In chapter two, after introducing the concept and application of geographic information system (GIS), the exploratory spatial data analysis method based on GIS is explained in detail. In addition, we also focused on the principles and applications of spatial econometric models, social network analysis and logistic regression models. Finally, we introduced the data used in this paper, including official statistics, emerging big data, and national migrant population dynamic monitoring and survey data.

In chapter three, we studied the scale of population mobility in 290 prefecture-level cities in China since 2000 based on national statistics. In particular, we focused on the changes in population mobility after 2010, and depicted the spatial patterns of China's domestic population mobility in the past 20 years. In addition, we have developed a mixed model to capture spatial effects and explore their impacts on population mobility

and heterogeneity effects from three different aspects of urban development, namely, public policy, habitant environment and social economy. The results show that the spatial patterns of population mobility in China has changed significantly since 2010. Although economic differences between regions are an important factor in attracting population mobility, the impact of urban habitant environment seems to be more and more important after 2010. Public policies, such as population control policies, have a weak positive impact on population mobility in megacities. These findings contribute to our understanding of China's population mobility patterns and their interaction with urban development at the macro level, and provide valuable information for governments and planners while developing effective strategies to promote sustainable urban development.

In chapter four, with the help of Tencent location big data, we describe the population mobility network of prefecture-level cities in China during the Spring Festival in 2019, and analyze the characteristics of the population mobility network with social network analysis method. In addition, by using the mixed model, we focus on the impact of urban socio-economic development on population mobility. The results show that the population mobility network presents a diamond structure, the top of the diamond is located in the four major urban agglomerations in China, and the population mobility network has a typical "small world" characteristic. The results of the mixed model show that urban social economy has a strong attraction for population mobility. Specifically, the wage level of the city, the development of the tertiary industry and foreign investment have a significant role in promoting the population inflow. The research results not only expand the application of big data, but also have a deeper understanding of the relationship between population mobility network and urban development.

In chapter five, we empirically analyze the impact of population inflow on regional and urban development at the city scale. In particular, we focus on China's Yangtze River Delta region and explore how population mobility affects urban economic development. The results of the panel model show that the urban GDP increases by 0.259% for every 1% increase of population inflow in the Yangtze River Delta urban agglomeration. Our results also show that the transmission mechanism of population inflow affects urban economy. Urban population inflow can transform industrial structure, promote residents' consumption, improve scientific and technological innovation ability and increase labor market employment, so as to realize the rapid growth of urban economy. Our results confirm the interactive relationship between population mobility and urban development, and supplement the research on urban population.

In chapter six, we use Baidu heat map data as the data support to analyze the dynamic changes of urban population mobility at the street block scale, and pay special attention to the impact of urban built environment on urban population mobility. In particular, on weekdays, the urban population presents the characteristics of central aggregation in the daytime and diffusion in the nighttime, while on the weekends, the opposite is true. The regression results show that the urban built environment quantified by 5Ds has a significant impact on the urban population mobility. The spatial form of buildings (Design), the diversity of urban functional facilities (diversity), the density of urban functional facilities (density), the intersection and density of urban road network (destination accessibility), and the transfer of regional center or public transport hub (distance to transit) can positively affect the flow of population in the city. Our results highlight the important role of the built environment in maintaining and increasing urban vitality. The flow of population in the city greatly promotes the prosperity of urban vitality, which is affected by the built environment of the city. Our results can provide inspiration for urban designers and urban planners, and provide theoretical support for them to plan a vibrant city.

In chapter seven, the whole summary of each chapter has been presented.

CONTENTS

Chapter 3. Spatiotemporal characteristics and driving forces of population mobility in urban China from 2000-2018

Chapter 4. Spatiotemporal patterns of population mobility and its determinants in urban China during Spring Festival of 2019

Chapter 5. Empirical Analysis on the Mechanism of Population Mobility Promoting Urban Development in China

Chapter 6. Spatiotemporal dynamics of population mobility and mechanism analysis at street block scale: A case study in Qingdao, China

Chapter 7. Conclusion

CONTENTS OF FIGURES

CONTENTS OF TABLES

Chapter 1. Background and purpose of the study

1.1 Introduction

Urban development is a complex system, initially proposed by demography experts, and urbanization is a direct external indicator to measure urban development. At present, from the global perspective, more people live in urban areas than in rural areas. In 2007, for the first time in history, the global urban population exceeded the global rural population, and the world population has remained predominantly urban thereafter (Fig. 1-1) [1]. The planet has gone through a process of rapid urbanization over the past six decades. In 1950, more than two-thirds (70 percent) of people worldwide lived in rural settlements and less than one-third (30 percent) in urban settlements. In 2014, 54 percent of the world's population is urban. The urban population is expected to continue to grow, so that by 2050, the world will be one-third rural (34 percent) and two-thirds urban (66 percent), roughly the reverse of the global rural-urban population distribution of the mid-twentieth century [2].

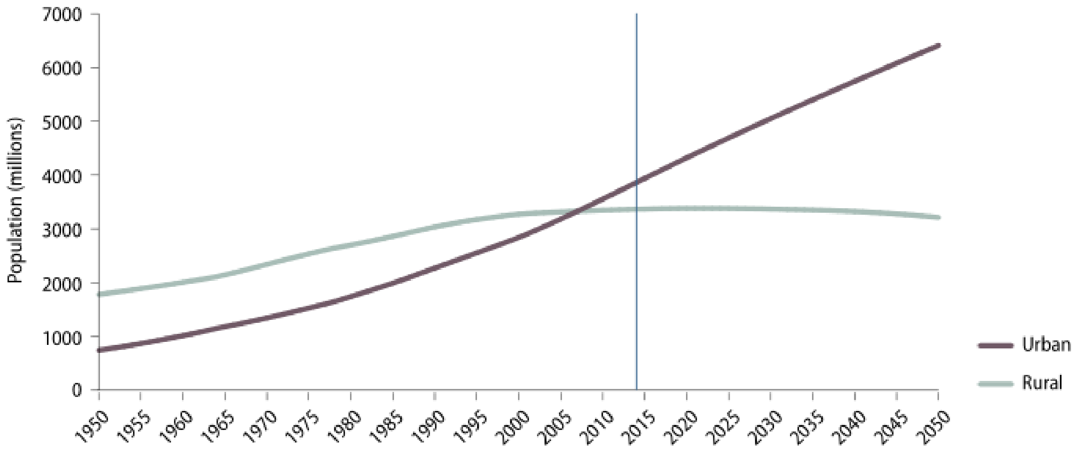


Fig. 1-1 Urban and rural population of the world, 1950-2050 (Source: UN World urbanization prospects: The 2018 revision).

However, levels of urbanization vary greatly across regions. In 2014, high levels of urbanization, at or above 80 percent, characterized Latin America and the Caribbean and North ern America. Europe, with 73 percent of its population living in urban areas, is expected to be over 80 percent urban by 2050 (Fig. 1-2). Africa and Asia, in contrast, remain mostly rural, with 40 percent and 48 percent of their respective populations living in urban areas. Over the coming decades, the level of urbanization is expected to increase in all regions (also referred to as major areas), with Africa and Asia urbanizing faster than the rest. Nevertheless, these two regions, which are projected to reach 56 and 64 percent urban by mid-century, respectively, are still expected to be less urbanized than other regions of the world. Africa and Asia are urbanizing more rapidly than other regions of the world. The rate of urbanization, measured as the average annual rate of change of the percentage urban, is highest in Asia and Africa, where currently the proportion urban is increasing by 1.5 and 1.1 percent per annum, respectively. Regions that already have relatively high levels of urbanization are urbanizing at a slower pace, at less than 0.4 percent annually (Fig. 4). In general, the pace of urbanization tends to slow down as a population becomes more urbanized [1].

China, as the largest country in Asia and the largest developing country in the world, has attracted great attention in its domestic urban development. Before 1978, the process of urbanization in China was quite slow and even declined. Between 1950 and 1980, the urbanization rate of developing countries rose from 16.2 percent to 30.5 percent, while that of China only rose from 11.2 percent to 19.4 percent. After the reform and opening up in 1978, the national economy developed rapidly, and the influx of labor force made the urban population surge, thus accelerating the process of urbanization. Since the beginning of the 21st century, China's urbanization rate has grown at an average annual rate of over 1%, much higher than the Asian average (Fig. 1-3 and Fig. 1-4).

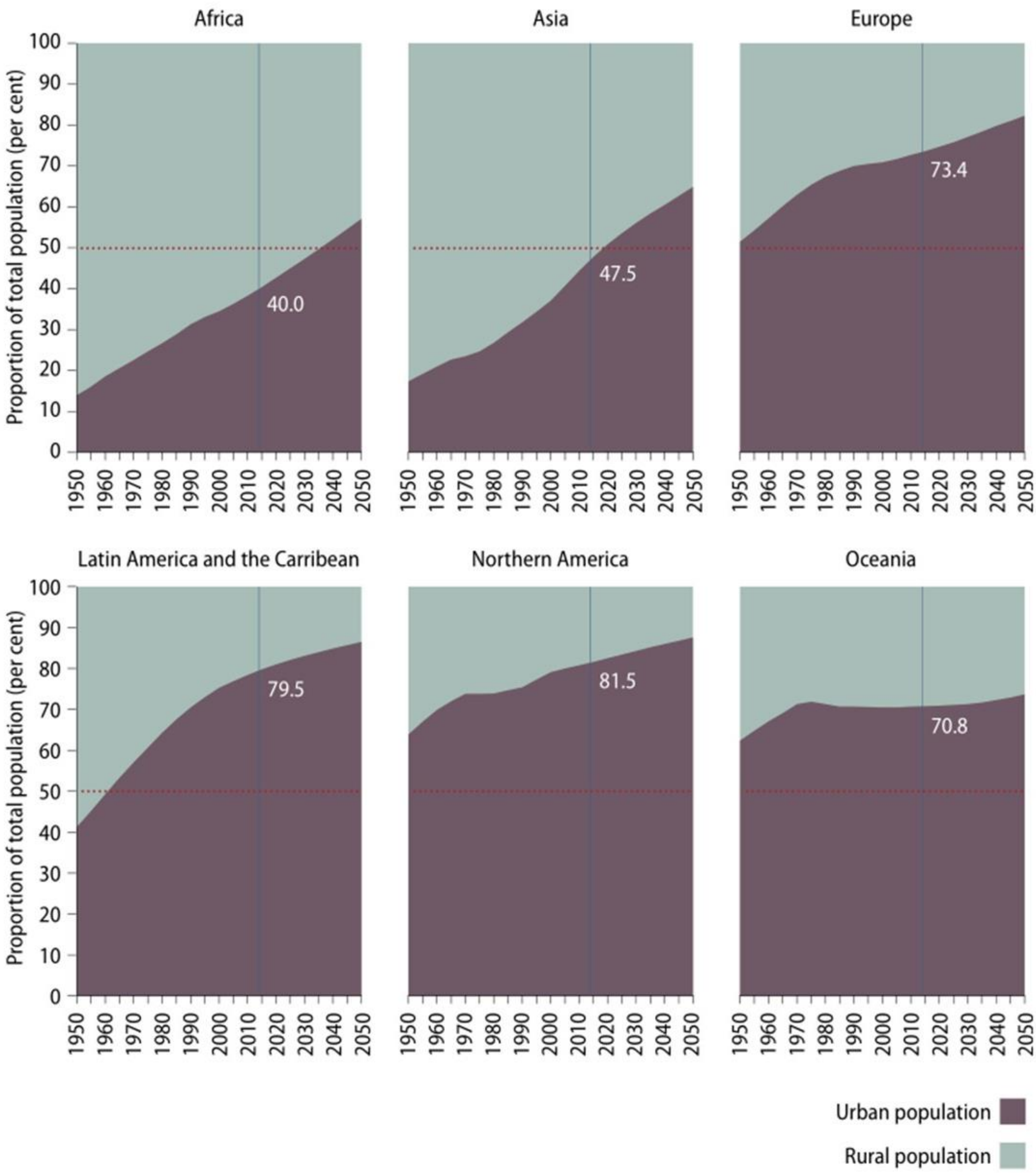


Fig. 1-2 Urban and rural population as proportion of total population
(Source: UN World urbanization prospects: The 2018 revision).

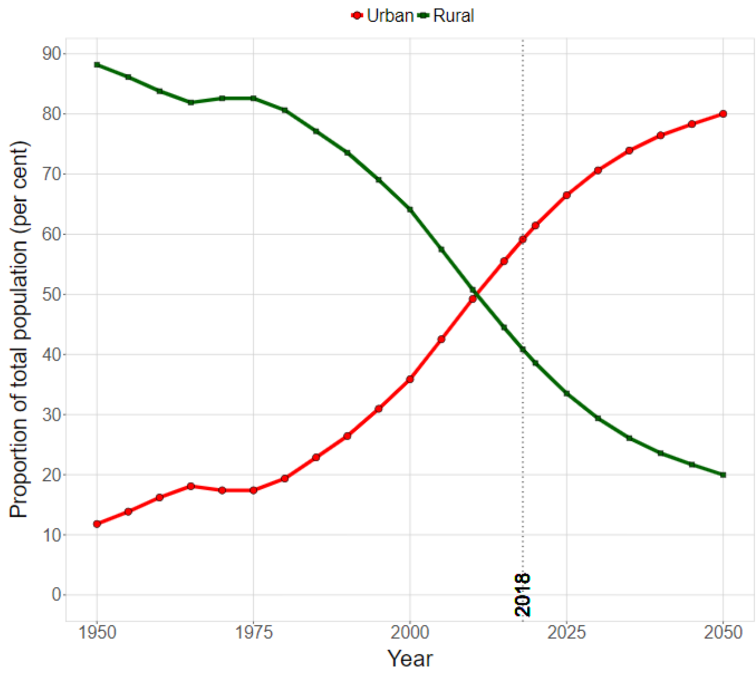


Fig. 1-3 Percentage of population in urban and rural areas in China (Source: UN World urbanization prospects: The 2018 revision).

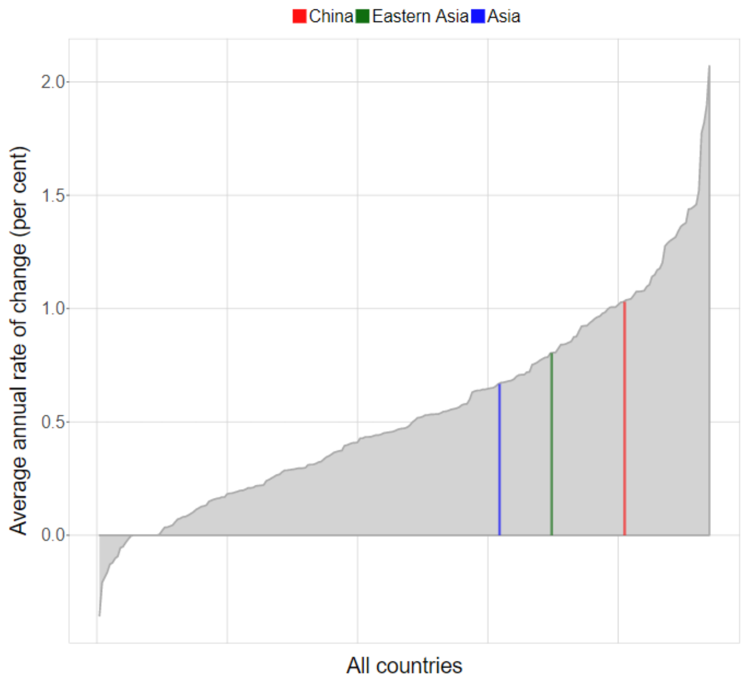


Fig. 1-4 Growth rate of the percentage urban, 1950-2018 (Source: UN World urbanization prospects: The 2018 revision).

Urbanization has also led to significant changes in the number of cities in China. In 1990, China had no megacities with a population of more than 10 million, and the number of large, medium-sized, and small cities was 2,32, and 37, respectively. However, by 2018, China has six cities with a population of more than 10 million, and the number of large cities, medium-sized cities and small cities has reached 13,105 and 160 respectively. It can be predicted that whether from the perspective of urbanization rate or the number of cities, there is still a huge space for China's urban development in the future, which will also become the most noticeable problem (Fig. 1-5) [3, 4].

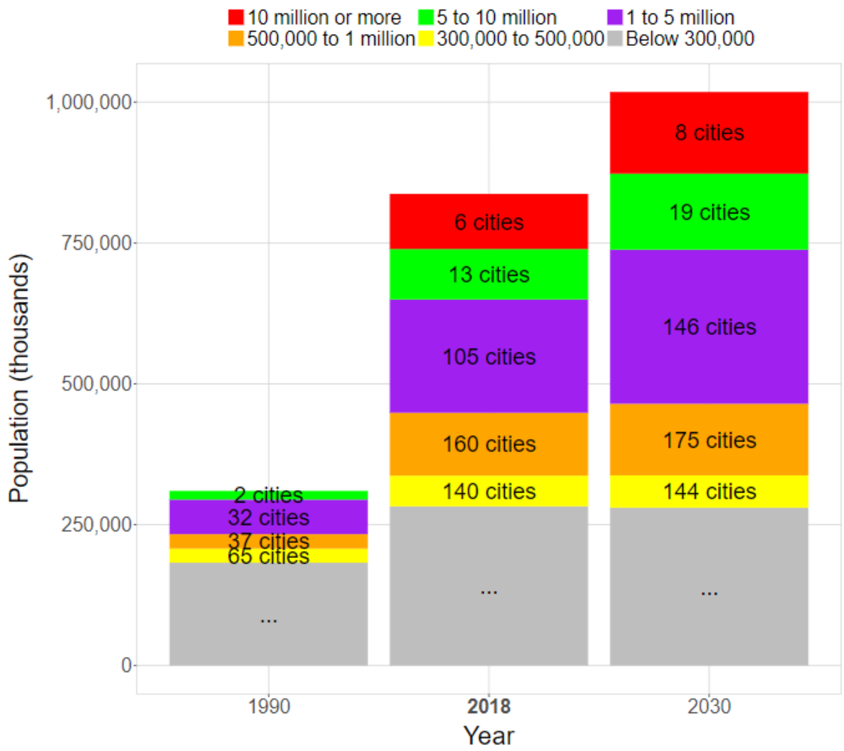


Fig. 1-5 Urban population by size class of urban settlement in China (Source: UN World urbanization prospects: The 2012 revision).

Different from western countries such as Europe and the United States, China's urbanization process has its own uniqueness. In the past 40 years, the flow of population from rural to urban areas and between regions has provided continuous impetus for the progress of China's economic and social development and the process of urbanization. Hundreds of millions of floating populations have realized China's urbanization through the way of "population flow". It can be said that China's "economic miracle" has been supported by large-scale population mobility and aggregation in the past (Fig. 1-6, 1-7) [5].

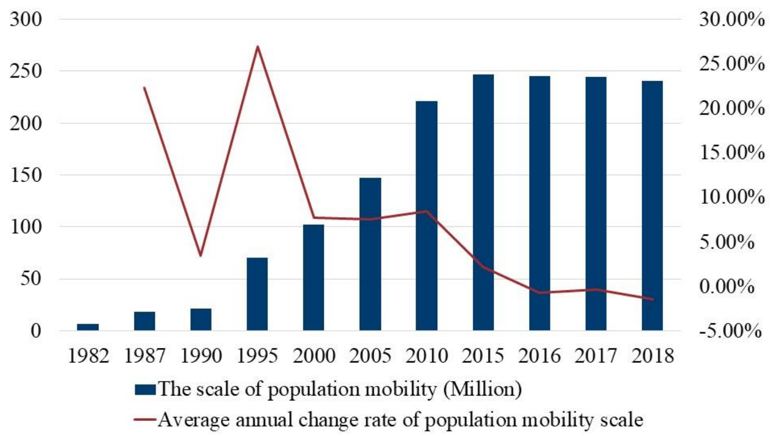


Fig. 1-6 The scale and average annual change rate of population mobility in China.

Population is the main body of economic activities, and labor force is an important element of economic development. As a barometer of the rise and fall of cities, population is not only the builders, producers and consumers of cities, but also the creators of new knowledge, the disseminators of new ideas and the applicators of new technologies. As the core of urban development, population mobility and aggregation have a significant impact on national economic geography, urban morphology and production mode, which is also the focus of urban economics and urban geography research [6, 7]. Cross regional population flow provides sufficient labor force for the development of the city, which is directly reflected in the rapid development of urban economy. At the same time, the rapid development of cities has become the "thrust" of population inflow, which continues to attract the input of labor force. Therefore, it is of great strategic significance to study the interactive relationship between population mobility and urban development for promoting the current urbanization process and maintaining the balanced development among regions.

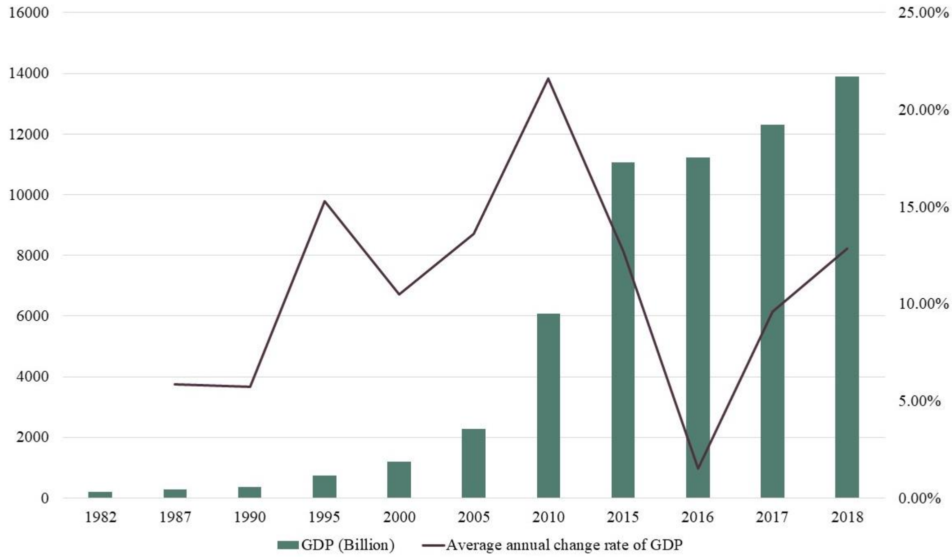


Fig. 1-7 The volume and average annual change rate of GDP in China.

1.2 Research background and significance

1.2.1 The complexity and diversity of urban development in China in the new era

Since the founding of the People's Republic of China, especially since the reform and opening up, China has experienced the fastest economic development and the most profound changes in its history, among which the large-scale mobility of population and labor force is one of the most important changes. At present, population mobility has become an important issue in regional economy and urban development, and the large-scale population mobility in China has become the focus of attention.

People are increasingly aware of the close relationship between population mobility and urban development. The expansion of population mobility has not only improved the regional economy, but also accelerated the urbanization process. However, different from developed countries, population mobility in China is restricted by multiple factors such as economic level, social structure and household registration system, showing certain complexity [8]. In addition, foreign views on population mobility and urban development are mainly used to explain the relationship between spontaneous population migration and economic and social development under the condition of full market economy, which cannot be fully applied to the conditions in China.

People oriented new urbanization has become the focus of national strategy. In the previous paper, based on historical data, we found that compared with developed countries in the world, China's urbanization level is not only low, but also low quality. In 2014, China's urbanization level just reached the world average level, while the urbanization rates of Japan, Britain, South Korea, the United States, France and Germany were 93%, 82%, 82%, 81%, 79% and 75%, respectively [2]. In addition, due to the restrictions of China's unique household registration system, many floating populations cannot enjoy the same treatment as urban residents, thus forming the "semi-urbanization" [9]. If the urbanization rate is counted according to the urban household registration population, in fact, China's urbanization level is only about 40%, which is much lower than that of developed countries [10]. It can be seen that China's urbanization level still has a large room for improvement in the future. To improve the quantity and quality of urbanization in the future will be a major practical problem faced by urban development, and it will also be the focus of multiple concerns.

In the 21st century, the macro background of urban development in China has changed, and the domestic and international economic situation has created new requirements for urban development. At the same time, it is also urgent to solve many social, environmental and ecological problems arising in the process of urbanization. Although the new urbanization emphasizes people-centered, insufficient attention has been paid to the interaction between population mobility and urbanization. Previous studies mainly focused on population agglomeration to cities, but ignored the location of agglomeration, which is of great significance to urban development, especially the urbanization process. Urban development and population mobility as a multi-dimensional complex social space process, it is necessary to carry out in-depth research from the perspective of multiple levels and two-way interaction.

1.2.2 Significant changes have taken place in population mobility in China

Since the 1980s, China's population mobility policy has gone through three stages. The first stage, from 1984 to the beginning of the 21st century, belongs to the stage of gradual opening up. Since 1984, farmers were allowed to work and do business in towns and market towns below the county level. The reform of citizenship management and food supply system also facilitated the flow of population. After the 1990s, the number of floating populations increased rapidly. At this stage, the policy of population mobility has been relaxed to a certain extent, but the degree is very limited, and it is mainly promoted by the central government from the top to the bottom, and the enthusiasm of local governments to support population mobility is not high [11].

From 2003 to 2012, it belongs to the stage of fair treatment of population mobility. In the 21st century, the concept of fair treatment of floating population has been put forward and implemented in this period, and a series of important changes have taken place in relevant policies. In 2006, the State Council issued the first systematic document on the issue of migrant workers, "some opinions on solving the problem of migrant workers", which put forward the basic principle of "fair treatment and equal treatment". In May 2012, the State Council issued the "12th Five Year Plan" of the national basic public service system, which provided a reliable system guarantee for the floating population to enjoy equal basic public services.

Since 2012, it belongs to the third stage, which is the stage of comprehensively promoting citizenization. The government report put forward "accelerating the reform of registered residence system, promoting the citizenization of agricultural transfer population in an orderly way, and striving to achieve the full coverage of the permanent resident population of urban basic public services". In March 2014, the CPC Central Committee and the State Council issued the national new urbanization plan (2014-2020). Subsequently, the State Council issued the "opinions on further improving the service for migrant workers" and "opinions on further promoting the reform of registered residence system". The policy framework for solving the floating population problem is clearer, and the sense of gaining and happiness of floating population is growing. The report of the 19th National Congress of the Communist Party of China further emphasizes breaking down the barriers hindering the flow of population and promoting the development of citizenization. It particularly requires that "the drawbacks of the system and mechanism hindering the social flow of labor force and talents be removed, so that everyone has the opportunity to realize their own development through hard work".

At the same time, the current status of population mobility in China has also changed significantly. Since the 1980s, the change process of the scale of population mobility in China can be roughly divided into three periods: The first period was from the early 1980s to the early 1990s. With the release of the notice on the issue of farmers entering market towns and settling down, the state relaxed the restrictions on the rural population entering small and medium-sized cities and towns to promote the rural population's rural-urban transfer. China's population mobility scale increased from 6.57 million in 1982 to 21.35 million in 1990, with an average annual growth of about 7%. In the second stage, from 1990 to 2010, the scale of population mobility increased at a faster speed, from 21.35 million in 1990 to 221.43 million in 2010, with an average annual growth of about 12%. The third stage is from 2010 to now, which is relatively mild. The growth rate of population mobility scale from 2010 to 2015 has decreased significantly, with an average annual growth of about 2%. Since 2015, new changes have taken place in the development of the population mobility. In 2015, the National Bureau of Statistics announced that the scale of population mobility in China was 247 million, down about 6 million compared with 2014; in 2016, the scale of floating population decreased by 1.71 million compared with 2015, and continued to decrease by 820000 in 2017.

In addition to the significant changes in the policy and scale of population mobility, the problems hidden behind the population mobility began to surface. Among them, the elderly population flow, children population flow and left behind children become the focus of attention. Among the population mobility, the scale of the elderly floating population has increased rapidly since 2000, from 5.03 million in 2000 to 13.04 million in 2015, with an average annual growth of 6.6%. The proportion of middle-aged and elderly floating population in China increased slightly from 2000 to 2015, 4.9% in 2000 and 5.3% in 2015.

Similar to the growth of floating population, the scale of migrant children in China has also experienced a transition from rapid growth to steady decline. From the overall scale, the total number of migrant children increased from 2.54 million in 1982 to 35.81 million in 2010, and then decreased. In 2015, the total number of migrant children was 34.26 million, which was 1.55 million lowers than that in 2010, with a decrease rate of about 4%. This is in contrast with the change direction of the national floating population, which increased by nearly 12% from 221 million to 247 million from 2010 to 2015. According to the sample data of the sixth population census of China in 2010, there are 61.255 million left behind children in rural areas, accounting for 37.7% of rural children and 21.88% of national children. Since 2010, the state has successively introduced a series of policies to solve the problems of left behind children in the population mobility. As of 2018, the number of left behind children in China is still close to 7 million (Fig. 1-8).

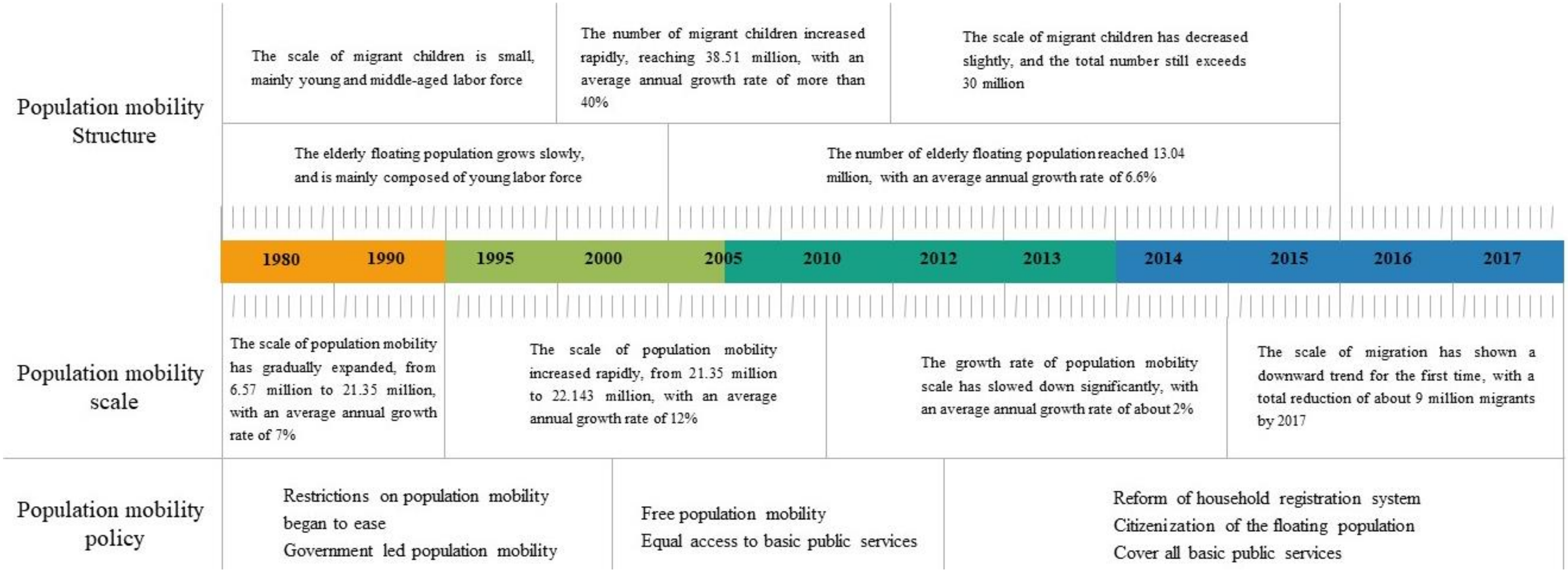


Fig. 1-8 The changing process of population mobility structure, scale and policy.

1.2.3 The potential of emerging big data in spatiotemporal behavior research

Population census and 1% population sampling survey are the most important data sources in the study of population mobility [12]. However, the data format restricts the spatial and temporal accuracy of the study. In the time dimension, population mobility is a large-scale social behavior with significant diversity in the process of time and space. The format of census data leads to the limitation of two-dimensional static perspective of related research. Due to the lack of a relatively complete data of population mobility in a relatively complete time, the current research results in the average situation of population spatial distribution pattern in different stages, which further leads to the fact that the specific links, evolution rules, spatial effects and influencing factors in the process of China's population mobility pattern have not been fully revealed for a long time. On the other hand, the 10-year census also has obvious time lag under the background of rapid urbanization in China.

In the spatial dimension, the accuracy of the research is improved through sampling survey and questionnaire survey in the meso scale [13]. The decision-making factors, social effects and spatial differences of micro themes are discussed in depth. However, the sampling survey brings fine research, but there are also some problems, such as the selection of research object is one-sided, regional aggregation and so on. Therefore, this kind of research is enough as an empirical analysis of local areas, and has insufficient explanatory power to the macro characteristics. At the macro scale, previous research conclusions such as: China's floating population shows the obvious characteristics of flowing from Anhui, Henan, Sichuan, Hunan, Guangxi and other central and western regions to the eastern coastal areas, and has formed the Yangtze River Delta, the Pearl River Delta, Beijing Tianjin Hebei and other highly concentrated floating population areas. This is not only an academic consensus, but also a common-sense cognition of the society [14, 15]. However, from the perspective of the spatial characteristics of China's population flow, the key points of the study are how the floating population changes in the process of aggregation and diffusion in the places of immigration and emigration, and the relationship between the them.

The emergence of emerging big data makes it possible to obtain human movement patterns from massive spatiotemporal trajectories based on individual granularity. The development of GPS, LBS and other technologies provides technical support for the observation of spatiotemporal characteristics of large-scale population behavior. Although there are significant uncertainties in human social activities, the data based on large-scale, large sample, dynamic space-time and relational positioning make the huge amount of human activities have rules to follow at the overall level [16, 17]. The big data with individual granularity spatiotemporal markers (including mobile phone call data, bus card data, social networking site check-in data, taxi track data, etc.) and the highly accurate spatiotemporal information of people, time and place in it provide sufficient research samples and precision for the research on population mobility. Based on this, this paper will try to combine the new big data to study the interaction between population mobility and urban development. Figs. 1-9 and 1-10 show the cross regional trajectory of population and the urban

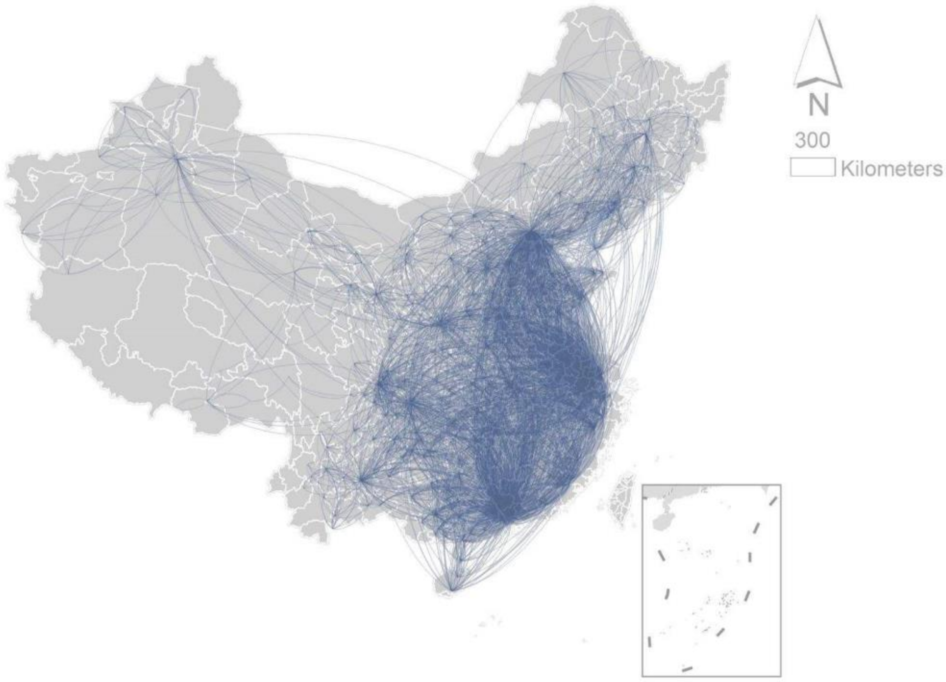


Fig. 1-9 Trajectory map of population mobility across regions based on mobile phone signaling data.

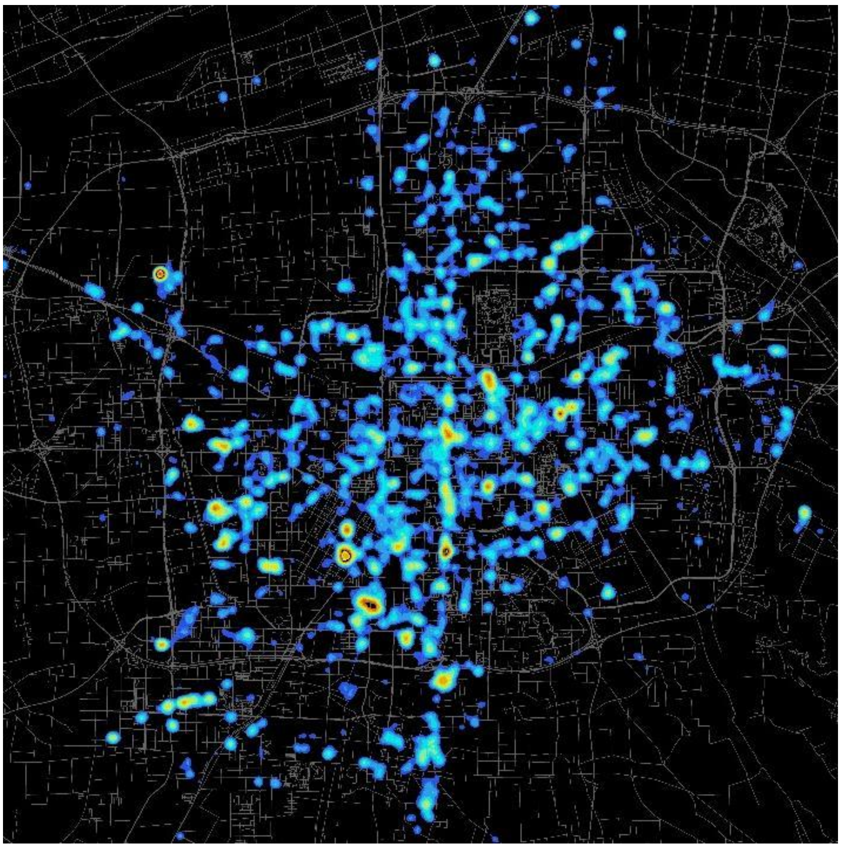


Fig. 1-10 Real time distribution map of urban population based on location service.

population distribution map at a specific time based on mobile phone data.

1.2.4 Research significance

Urban development usually refers to the change and growth process of the status, role, attraction and radiation force of a city in a certain region, and it is also a process to meet the multi-level population growth. Generally speaking, the progress of economic level reflects the growth of "quantity" of cities, while the level of urbanization reflects the improvement of "quality" of cities. Previous studies also measure the development of a city from its economic level and urbanization level. However, the development of the city is a complex dynamic process, both external development and internal development. Although economy and urbanization can represent urban development to a certain extent, they are still not comprehensive enough. The progress of urban level should be reflected in the "soft power" of the city, such as the supply of public services and the planning and construction of the city. In this paper, we build the multi-scale analysis model, respectively from the macro, meso and micro perspective to investigate the population mobility and interactive relationship of urban development, this aspect can start from the external environment for the development of the city, especially urbanization provides the result of empirical analysis, at the same time, emphasize the inner cities "soft power" on the appeal of the important role that highlights the importance of the city in terms of population mobility.

As we all know, there is a two-way relationship between population mobility and urban development. The large-scale inflow of population promotes the growth of regional economy and makes cities develop rapidly. At the same time, the developed cities become the destination of other floating population. In the long run, the relationship between population mobility and urban development has become more and more complex. Previous studies only based on macro statistical data, or investigate the impact of population mobility on urban development, or analysis of the attractiveness of urban development to population mobility, but lack of systematic and comprehensive understanding. This paper examines the two-way impact of population mobility and urban development, and focuses on the impact of urban development on population mobility, aiming to provide some data reference for the development of cities in that the population mobility scale decline.

At the same time, compared with the previous official data channels such as census, statistical yearbook and sampling survey, this paper attempts to use the emerging big data to study the spatial pattern of population mobility in China, which has expanded the previous research perspective and has certain practical value. we used the spatial econometric model to analyze the interaction between population mobility and urban development after considering the spatial effect of spatial data, which provides a new way for the research of the current urbanization process. Urban development, urbanization and population mobility are the key research topics in urban geography, demography, urban economics and other disciplines. For China, which is transformed under the background of globalization, information and rapid urbanization, the spatial pattern change of floating population is not only shaping the spatial pattern of Chinese population for a long time, but also has a significant impact on the process of urbanization in China. It is of great practical significance to study the special phenomenon of population flow by combining various data and systematically for understanding the evolution trend of urbanization development in China and guiding the future urban construction and urban policy making. As an important part of global economic and regional development issues, population mobility has become the core issue of concern for scholars, policy makers and urban planners.

1.3 Review of previous study

1.3.1 Study on the urban development

As a separate object, urban development is mainly studied from the perspective of urbanization. As we all know, in the past few decades, the world is experiencing rapid urbanization, especially in China [18, 19]. With the continuous migration of urban and rural population and the continuous expansion of cities, more and more megacities and big cities have emerged in China [20, 21]. However, the urbanization level of different cities in China is quite different. With the rapid development of urbanization and modernization, urban built-up land has been replaced by natural landscape, resulting in a series of “urban diseases”, such as water pollution, air pollution, greenhouse gas increase, fossil fuel energy consumption increase, urban heat island and traffic congestion [22-24]. Therefore, urbanization does not appear alone, but is accompanied by a series of environmental problems from the perspective of urbanization [25-27]. With the close economic ties and transportation, regional studies based on urban agglomerations have gradually replaced single cities and become the focus of research, such as Beijing Tianjin Hebei urban agglomeration, Yangtze River Delta urban agglomeration and Pearl River Delta urban agglomeration [28].

As a complex urban problem, urbanization cannot be evaluated only by population [29]. Economic and social factors are the important contents of evaluating urbanization and quantifying urban development [30-33]. At present, many studies start from the evaluation system of regional urbanization, trying to use AHP, fuzzy comprehensive evaluation method, entropy method and gray correlation method to comprehensively compare the level of regional urbanization [34-37]. In the urbanization coupling coordination model, population, economic development and social development are regarded as three independent and integrated subsystems, which are nonlinear and interactive. They are widely used as a new standard to judge the level of urbanization [38].

1.3.2 Study on the population mobility and its relationship between urban development

In the research on the interaction between population mobility and urban development, many scholars have studied the attraction of economic development of destination cities to population mobility from the perspective of economic benefits. Many influential theories and methods have been proposed to explain the origin, mechanism and development of population migration. The representative theories include “immigration law”, “push-pull theory”, “macro-neoclassical theory”, “micro-neoclassical theory”, “equilibrium theory” and Lewis-Fei-Ranis model. Ravenstein’s immigration law first proposed civilized laws and regulations related to population migration, and systematically analyzed the phenomenon of population migration from the aspects of migration reasons, distance and characteristics [39]. Bogue [40] tries to use push-pull theory to explain the internal mechanism of population migration from the the aspects of convenient employment, familiar social network, high-risk unemployment rate and so on. Macro-neoclassical theory holds that the expected income level of the future destination is the main driving force of population migration [41, 42]. Micro-neoclassical theory abstracts population migration as a form of capital investment, which aims to maximize personal interests [43]. The equilibrium theory emphasizes that the amenities (environment, service and living) of the destination are more likely to have an impact on population migration than the income differences among regions [44, 45]. The structural approach that originated in the Lewis-Fei-Ranis model [41, 42] is another method to study population migration [46]. This approach emphasizes the importance of labor market segmentation, unbalanced economic opportunities and the background of immigration system. Compared with the previous migration theories, the structural approach highlights the influence of national institutional factors and emphasizes that state intervention has a great impact on internal migration [47, 48].

Recent studies have also highlighted the impact of urban amenities, public services, social integration and subjective feelings of migrants on regional migration [49-54]. The migration of “lifestyle” and “social relations” has been growing steadily, but in the case of little difference in economic level, family factors (such as family size), social factors (such as industry, social security) and social network factors have gradually become the focus of attention [55-57]. It can be seen that with the deepening of the research, in addition to the development of urban economy, the social development of the city also has a greater impact on the regional population mobility.

A large amount of papers has studied the phenomenon and mechanism of population migration in China since the reform and opening up. Due to China’s unique institutional factors and economic development process, macroeconomic development and national institutional intervention have become the two most in-depth and comprehensive research areas.

Economic factors also play an important role in promoting China’s population migration [58, 59]. The coastal development strategy, reform and opening policy allow some areas to be developed first, thus the regional development focus is shifted to the southeast coastal areas. After making full use of the advantages of location and technology, the rapid economic growth has appeared in the Pearl River Delta, Yangtze River Delta and Beijing-Tianjin-Hebei region, which makes the regional development gap between the eastern, central and western regions more and more large, so there is an obvious phenomenon of inter-provincial migration [60-63]. The continuous improvement of economic conditions has brought more opportunities for immigrants in the above-mentioned areas and has become the preferred destination for population migration [64]. After the 21st century, with a series of problems, such as economic transformation, industrial upgrading and investment costs, labor-intensive industries began to transfer to the central and western regions, which to a certain extent attracted the population migration, while the eastern coastal areas, relying on the high-end industries and service industries, has an advantage in attracting technical talents [65].

However, with the reduction of population mobility scale, the change of population mobility structure and various problems in the process of urbanization, regional economy has no longer become the only factor to attract population inflow. Recent studies have pointed out that changes in inhabitant environment can also affect population migration between regions. Lin, Zhu, Ke and Wang [54] explored the impact of urban basic public services on population migration in eastern 6 provinces based on micro-survey data. Jianan [52] and Wang [50] emphasized that the social integration of migration destinations is an important factor affecting population migration. In addition, individual factors, family factors and social network factors can affect the regional population migration [55-57, 66]. For the first and second tier cities, housing prices, air quality and traffic conditions are increasingly becoming the important indicators of their immigration destinations [67]. However, we still don't know the changes of these factors in spatial and temporal dimension.

In addition to inter-city and inter-regional population mobility, intra-city population mobility is also noteworthy. The population mobility between different regions within a city usually reflects the urban vitality. With the emergence of emerging big data, urban vitality characterized by urban residents' activity trajectory has become the focus of research. Mobile phone data, location-based service (LBS) data, social media data, and geotagged data, have provided significant advantages, allowing us to capture the dynamic interaction between residents and urban spaces. The research based on the above data sources have been widely conducted in urban space, like Long and Zhou [68] used mobile phone data, Li, et al. [69] used mobile phone positioning records, Wu, et al. [70] combined GPS tracking data, Zheng, et al. [71] adopted catering establishment data and Shen and Karimi [72] used housing price records. These new datasets can generate specific images of urban vitality, allowing us to explore the characteristics and internal mechanism of urban vitality from many aspects.

The multi-source data mining, which reflects the trajectory of population mobility and population information, provides data support for us to study the interaction between population mobility and urban development at different scales. Data sources with different precision are selected in different research scales. For example, at the national scale, the support of national statistical yearbook and Tencent location big data provide guarantee for us to study population mobility and urban development at a macro scale. At the local level, with the help of the heat map data reflecting the real-time population distribution and the questionnaire data for the floating population, we can explore the relationship between population mobility and urban development from the perspective of urban planning design and urban development policies.

1.4 Purpose and content overview of this study

This study focuses on the interactive relationship between population mobility and urban development in China. In particular, we not only focus on population mobility between cities, but also use emerging big data to depict the current situation of population mobility within cities. On this basis, the bidirectional relationship between population mobility and urban development is studied on different scales. We are particularly concerned about the changing attractiveness of urban development to mobility, which is crucial to China's future urbanization process. At the same time, we also emphasize how urban planning and design can influence population mobility in the region in order to guide the healthy development of cities. The chapter names and basic structure of this study are shown in Fig. 1-11. The brief chapters are shown in Fig. 1-12.

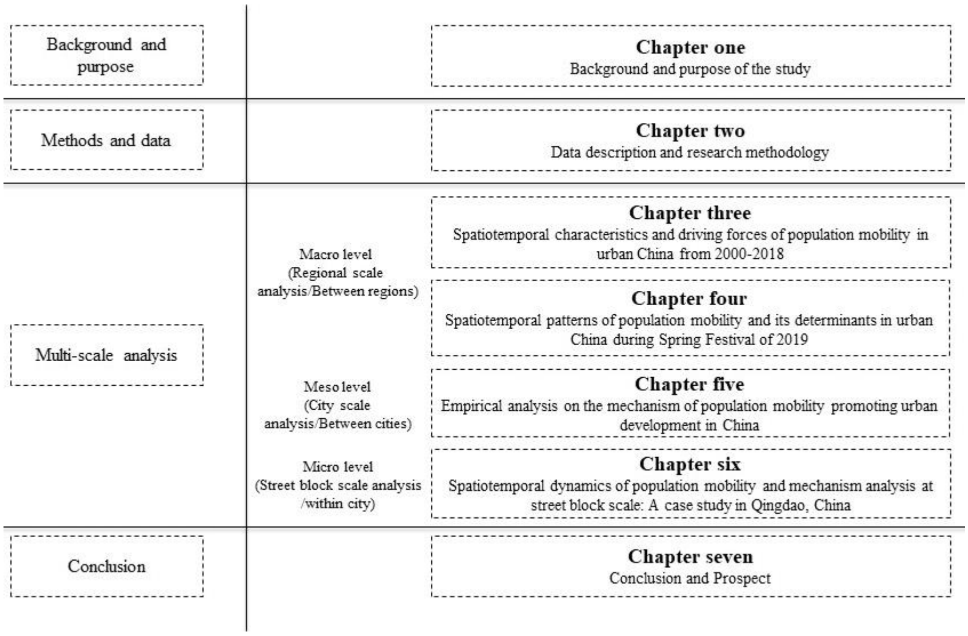


Fig. 1-11 Chapter name and basic structure.

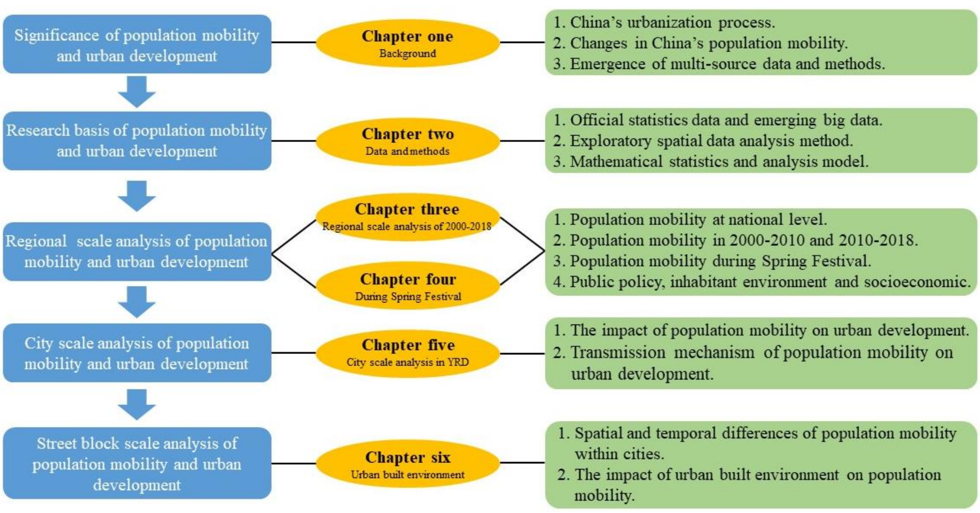


Fig. 1-12 Brief chapter introduction.

In Chapter 1, research background and purpose of the study:

Since the reform and opening up, China's population mobility has greatly promoted the development of regions and cities. However, with the change of China's population structure, the prominent problems in the process of urbanization and the change of population mobility, China's demographic bonus is gradually decreasing. Therefore, we need a comprehensive and systematic description of the interaction between population mobility and urban development. Previous studies have suggested that large-scale population mobility has changed regional economy and promoted urban prosperity. However, previous studies are more based on unilateral analysis. Population mobility and urban development are a process of mutual influence. Only considering the single link in the analysis will inevitably make the analysis incomplete. At the same time, the correction of population mobility within the city should be concerned, because it can put forward new suggestions from the perspective of urban planning. Therefore, after introduced the research background, we reviewed the previous research literature. Finally, we proposed the purpose of this study and expounded the research content.

In Chapter 2, research methods and data description:

Firstly, the concept and application of GIS was introduced. Secondly, we introduced the spatial analysis method based on GIS, that is exploratory spatial data analysis method. Because of the spatial effect of population mobility and urban development, we used spatial econometric model to explore the urban development factors that affect population mobility. In particular, we also introduced the social network analysis method and the logistic regression model, which are used to identify the spatial and temporal patterns of population mobility, and the empirical analysis based on the survey data. Finally, we introduced the data used in this paper in detail, including official statistics data, Tencent location big data, Baidu heat map data and national migrant population dynamic monitoring survey data.

In Chapter 3, spatiotemporal patterns and driving forces of population mobility in China during 2000-2018:

In the third chapter, we systematically summarize and sort out the population mobility in China in the past, especially supplement the changes of population mobility pattern after 2010. With the help of national official statistics, we describe the change characteristics of China's cross regional population mobility from 2000 to 2010 and from 2010 to 2018. At the same time, we empirically analyze the impact of urban development on population mobility and its change process from three different aspects of urban development, namely, public policy, social economy and inhabitant environment. Our results show that China's cross regional population mobility presents different spatial and temporal patterns in the above two periods. Although the scale of population mobility in the eastern region is still the largest, with the rise of the central and western regions and industrial migration, the growth rate of the increase shows a significant downward trend. At the same time, the provincial capital cities and the first and second tier cities in the central and western regions have become new destinations for population mobility. Our results also show that the impact of urban development on population mobility is different in different time periods. Before 2010, the economic level of cities was the main driving force to attract population inflow. However, with the change of population structure in China, the impact of public policies and inhabitant settlements on population mobility seems to be greater.

In Chapter 4, spatiotemporal patterns and determinants of population mobility in China during Spring Festival of 2019:

In this chapter, we try to describe the spatiotemporal network of population mobility by combining Tencent location big data with high spatial and temporal resolution. Specifically, compared with the official statistics, Tencent location big data not only ensures high accuracy in time and space, but also carries the temporal and spatial attributes of the target, which can more clearly describe the network structure of population mobility. At the same time, we limit the research time to the Spring Festival of 2019, which is the most appropriate time to study population mobility, because before and after Spring in China, a large number of people go back and forth between work place and home. With the help of location big data, this phenomenon can be easily captured. The analysis results show that the population mobility presents a spatial structure of “two horizontals and three verticals”. The major cities in the four urban agglomerations in China occupy an absolute core position in the population mobility network hierarchy, and the population mobility forms 11 different community structures. regression model results show that population mobility is significantly related to the value-added of secondary and tertiary industries, foreign capital, average wage, urbanization rate and value-added of primary industries. when the spatial heterogeneity and non-stationarity was considered, the socio-economic factors that affect population mobility exhibited differentiation among different regions and cities. With the help of emerging big data, we successfully describe the temporal and spatial pattern of population mobility in a short period of time, making up for the limitations of official statistical data analysis. At the same time, we also empirically analyze the impact of urban development, especially urban economic development, on population mobility with the help of spatial econometric model.

In Chapter 5, empirical analysis on the mechanism of population mobility promoting urban development in China—Case study in Yangtze River Delta urban agglomeration:

In the first two chapters, we used different data sources to describe the spatial-temporal patterns of population mobility in China at the regional scale, and analyze the impact of urban development on population mobility. However, there is a certain interaction between population mobility and urban development. While urban development affects population mobility, the influx of labor forces promotes the development of regional economy. Therefore, in this chapter, we study the impact of population mobility on urban development from the city scale. In order to express the urban development intuitively, we use the GDP of the city as the index to measure the urban development, and take the Yangtze River Delta urban agglomeration as the research object. Our results show that population mobility has a significant role in promoting regional economy. Specifically, for every 1% increase in population mobility scale, regional GDP will increase by 0.259%. We also found that regional population mobility promotes economic growth by transforming industrial structure, increasing residents' consumption, increasing technological innovation and improving employment. Therefore, in this part, we empirically analyze the impact of population mobility on urban development and its transmission mechanism, which provides a supplement for the study of the interaction between them.

In Chapter 6, Spatiotemporal dynamics of population mobility and mechanism analysis at street block level:

The level of urban economy as a macro performance of urban development is discussed in the previous chapters. To some extent, urban design can also be regarded as the embodiment of urban development. Therefore, in this chapter, we focus on the street block level analysis of how urban design affects urban population mobility. We empirically analyze the impact of urban built environment (design, density, diversity, transfer distance and destination accessibility) on population mobility by combining Baidu heat map data and spatial econometric model that reflecting the change of urban population real-time distribution. Our study found that there were significant differences in the patterns of population mobility in working days, rest days, daytime and nighttime, and showed a polycentric distribution. To some extent, this reflects the characteristics of urban residents' daily commuting. It is important that the urban built environment has a significant impact on the population mobility within the city. Specifically, the morphology of buildings, the density of road intersections, the density of urban function and the accessibility of road network can improve the population flow in the region. These analyses can help urban decision makers and planners to optimize the urban spatial layout, which has important practical significance for high-quality urban design and social welfare.

In Chapter 7, conclusion and prospects:

This part summarized the research of previous chapters. And based on the conclusions, the future research of population mobility and urban development as well as the prospect of further research are put forward.

Reference:

1. Nations, U., Revision of World Urbanization Prospects. *United Nations: New York, NY, USA* **2018**.
2. Nations, U., World Urbanization Prospects: The 2014 Revision.: United Nations. New York. In 2018.
3. Habitat, U., *State of the world's cities 2012/2013: Prosperity of cities*. Routledge: 2013.
4. Nations, U., The World’ s Cities in 2018 - Data Booklet. **2018**.
5. Zhang, K. H.; Shunfeng, S., Rural–urban migration and urbanization in China: Evidence from time-series and cross-section analyses. *China Economic Review* **2003,** 14, (4), 386-400.
6. Chan, K. W., China: internal migration. *The encyclopedia of global human migration* **2013**.
7. Liang, Z., The age of migration in China. *Population and development review* **2001,** 27, (3), 499-524.
8. Yang, Z., Floating population and urbanization. *Population and economy* **1996,** 05, 33-38.
9. Wang, C., The Problem of Semi-Urbanization of Rural Floating Population. *Sociological research* **2006,** 5, 107-123.
10. Li, K., Deeply implement the strategy of expanding domestic demand in the process of reform and opening up. *Seeking truth* **2012,** 4, 3-10.
11. Commission, N. H., China’s Migrant Population Development Report. In China Population Publishing House: Beijing, China: 2018.
12. Li, S.-M., Population migration and urbanization in China: A comparative analysis of the 1990 population census and the 1995 national one percent sample population survey. *International Migration Review* **2004,** 38, (2), 655-685.
13. Tian, M., Spatial moblity of agricultural transfer population and urban integration. *Population research* **2013,** 37, (4), 43-55.
14. Tong, Y.; Wang, Y., The Choice of China's Floating Population: Why Beijing, Shanghai and Guangzhou are so popular——An analysis based on individual costs and benefits. *Population research* **2014,** 39, (4), 49-56.
15. Zhang, Y.; Cen, Q., China's population spatial mobility pattern and analysis of the influencing factors of interprovincial mobility. *Population research* **2014,** 38, (5), 54-71.
16. Long, Y.; Liu, L., Transformations of urban studies and planning in the big/open data era: A review. *International Journal of Image and Data Fusion* **2016,** 7, (4), 295-308.
17. Yang, Z.; Long, Y., Opportunities and limitations of big data applications to human and economic geography: The state of the art. *Progress in Geography* **2015,** 34, (4), 410-417.
18. Klugman, J., *A Sourcebook for Poverty Reduction Strategies: Volume 2: Macroeconomic and Sectoral Approaches*. World Bank Washington, DC: 2002.
19. Pallagst, K., The future of shrinking cities: problems, patterns and strategies of urban transformation in a global context. **2009**.
20. Chen, Q.; Song, Z., Accounting for China's urbanization. *China Economic Review* **2014,** 30, 485-494.
21. dos Santos, A. R.; de Oliveira, F. S.; da Silva, A. G.; Gleriani, J. M.; Gonçalves, W.; Moreira, G. L.; Silva, F. G.; Branco, E. R. F.; Moura, M. M.; da Silva, R. G., Spatial and temporal distribution of urban heat islands. *Science of the Total Environment* **2017,** 605, 946-956.
22. Foley, J. A.; DeFries, R.; Asner, G. P.; Barford, C.; Bonan, G.; Carpenter, S. R.; Chapin, F. S.; Coe, M. T.; Daily, G. C.; Gibbs, H. K., Global consequences of land use. *science* **2005,** 309,

(5734), 570-574.

23. Qin, B.; Zhang, Y., Note on urbanization in China: Urban definitions and census data. *China Economic Review* **2014,** 30, 495-502.
24. Zhou, L.; Dickinson, R. E.; Tian, Y.; Fang, J.; Li, Q.; Kaufmann, R. K.; Tucker, C. J.; Myneni, R. B., Evidence for a significant urbanization effect on climate in China. *Proceedings of the National Academy of Sciences* **2004,** 101, (26), 9540-9544.
25. Grimm, N. B.; Foster, D.; Groffman, P.; Grove, J. M.; Hopkinson, C. S.; Nadelhoffer, K. J.; Pataki, D. E.; Peters, D. P., The changing landscape: ecosystem responses to urbanization and pollution across climatic and societal gradients. *Frontiers in Ecology and the Environment* **2008,** 6, (5), 264-272.
26. Peng, S.; Piao, S.; Ciais, P.; Friedlingstein, P.; Ottle, C.; Bréon, F. o.-M.; Nan, H.; Zhou, L.; Myneni, R. B., Surface urban heat island across 419 global big cities. *Environmental science & technology* **2012,** 46, (2), 696-703.
27. Zhu, H.-M.; You, W.-H.; Zeng, Z.-f., Urbanization and CO2 emissions: A semi-parametric panel data analysis. *Economics Letters* **2012,** 117, (3), 848-850.
28. Zhou, D.; Zhao, S.; Zhang, L.; Liu, S., Remotely sensed assessment of urbanization effects on vegetation phenology in China's 32 major cities. *Remote Sensing of Environment* **2016,** 176, 272-281.
29. Zhu, X.; Gao, W.; Zhou, N.; Kammen, D. M.; Wu, Y.; Zhang, Y.; Chen, W., The inhabited environment, infrastructure development and advanced urbanization in China's Yangtze River Delta Region. *Environmental Research Letters* **2016,** 11, (12), 124020.
30. Chen, M.; Lu, D.; Zha, L., The comprehensive evaluation of China's urbanization and effects on resources and environment. *Journal of Geographical Sciences* **2010,** 20, (1), 17-30.
31. Chen, M.; Lu, D.; Zhang, H., Comprehensive evaluation and the driving factors of China's urbanization. *Acta Geographica Sinica* **2009,** 64, (4), 387-398.
32. Gallagher, R.; Appenzeller, T.; Normile, D., Beyond reductionism. *Science* **1999,** 284, (5411), 79.
33. Guo-ping, L., The analysis and evaluation of the compatible relationship between industrialization and urbanization in China. *Areal Research and Development* **2008,** 5, (002).
34. Beskese, A.; Demir, H. H.; Ozcan, H. K.; Okten, H. E., Landfill site selection using fuzzy AHP and fuzzy TOPSIS: a case study for Istanbul. *Environmental Earth Sciences* **2015,** 73, (7), 3513-3521.
35. Liu, Y.; Li, R.; Song, X., Grey associative analysis of regional urbanization and eco-environment coupling in China. *Acta Geographica Sinica* **2005,** 2, 237-247.
36. Sakieh, Y.; Amiri, B. J.; Danekar, A.; Feghhi, J.; Dezhkam, S., Scenario-based evaluation of urban development sustainability: an integrative modeling approach to compromise between urbanization suitability index and landscape pattern. *Environment, Development and Sustainability* **2015,** 17, (6), 1343-1365.
37. Wang, S.; Ma, H.; Zhao, Y., Exploring the relationship between urbanization and the eco-environment—A case study of Beijing–Tianjin–Hebei region. *Ecological Indicators* **2014,** 45, 171-183.
38. Yanbing, G., Comprehensive assessment on ecological risk of Hexi Corridor urbanization based on normal cloud model and entropy weight. *Journal of Arid Land Resources and Environment* **2012,** 26, (5), 169-174.

39. Ravenstein, E. G., The laws of migration. *Journal of the statistical society of London* **1885,** 48, (2), 167-235.
40. Bogue, D. J., The use of place-of-birth and duration-of-residence data for studying internal migration. *Ekistics* **1960,** 9, (52), 97-101.
41. Lewis, W. A., Economic development with unlimited supplies of labour. **1954**.
42. Ranis, G.; Fei, J. C., A theory of economic development. *The American economic review* **1961**, 533-565.
43. Sjaastad, L. A., The costs and returns of human migration. *Journal of political Economy* **1962,** 70, (5, Part 2), 80-93.
44. Knapp, T. A.; Gravest, P. E., On the role of amenities in models of migration and regional development. *Journal of regional science* **1989,** 29, (1), 71-87.
45. Mueser, P. R.; Graves, P. E., Examining the role of economic opportunity and amenities in explaining population redistribution. *Journal of Urban Economics* **1995,** 37, (2), 176-200.
46. Fan, C. C., Economic opportunities and internal migration: a case study of Guangdong Province, China. *The Professional Geographer* **1996,** 48, (1), 28-45.
47. Chan, K. W.; Zhang, L., The hukou system and rural-urban migration in China: Processes and changes. *The China Quarterly* **1999**, (160), 818-855.
48. Vendryes, T., Migration constraints and development: Hukou and capital accumulation in China. *China Economic Review* **2011,** 22, (4), 669-692.
49. Fan, C. C., Settlement intention and split households: Findings from a survey of migrants in Beijing's urban villages. *China Review* **2011**, 11-41.
50. Wang, Y., Settlement intention of rural migrants in Chinese cities: Findings from a twelve-city migrant survey. *Population research* **2013,** 37, (4), 19-32.
51. Cheng, Z.; Nielsen, I.; Smyth, R., Access to social insurance in urban China: A comparative study of rural–urban and urban–urban migrants in Beijing. *Habitat International* **2014,** 41, 243-252.
52. Jianan, Q., An empirical research of migrants' wish of leaving the city: The preliminary analysis based on a survey conducted in Beijing, Shanghai and Guangzhou. *Population Journal* **2014,** 36, (5), 80-86.
53. Liu, Y.; Xu, W., Destination choices of permanent and temporary migrants in China, 1985–2005. *Population, Space and Place* **2017,** 23, (1), e1963.
54. Lin, L.; Zhu, Y.; Ke, W.; Wang, j., The impact of migrants' access to urban public services on their urban settlement intentions: A study from the perspective of different-sized cities. *Acta Geographica Sinica* **2019,** 74, (4), 737-752.
55. Zhu, Y., China's floating population and their settlement intention in the cities: Beyond the Hukou reform. *Habitat International* **2007,** 31, (1), 65-76.
56. Hu, F.; Xu, Z.; Chen, Y., Circular migration, or permanent stay? Evidence from China's rural–urban migration. *China Economic Review* **2011,** 22, (1), 64-74.
57. Huang, X.; Liu, Y.; Xue, D.; Li, Z.; Shi, Z., The effects of social ties on rural-urban migrants' intention to settle in cities in China. *Cities* **2018,** 83, 203-212.
58. Fan, C. C., Modeling interprovincial migration in China, 1985-2000. *Eurasian Geography and Economics* **2005,** 46, (3), 165-184.
59. Shen, J., Increasing internal migration in China from 1985 to 2005: Institutional versus economic drivers. *Habitat International* **2013,** 39, 1-7.

60. Liang, Z., Foreign investment, economic growth, and temporary migration: The case of Shenzhen special economic zone, China. *Development and Society* **1999**.

61. Fang, C.; Dewen, W., Migration as marketization: what can we learn from China's 2000 census data? *China Review* **2003**, 73-93.

62. Zhu, Y., The floating population's household strategies and the role of migration in China's regional development and integration. *International Journal of Population Geography* **2003,** 9, (6), 485-502.

63. Fan, C. C.; Sun, M., Regional inequality in China, 1978-2006. *Eurasian geography and Economics* **2008,** 49, (1), 1-18.

64. Yeung, Y.-m.; Shen, J., Coastal China's urban-rural spatial restructuring under globalization. *Reshaping economic geography in East Asia* **2009**, 294-319.

65. He, C.; Chen, T.; Mao, X.; Zhou, Y., Economic transition, urbanization and population redistribution in China. *Habitat International* **2016**, (51), 39-47.

66. Liu, Z.; Wang, Y.; Chen, S., Does formal housing encourage settlement intention of rural migrants in Chinese cities? A structural equation model analysis. *Urban Studies* **2017,** 54, (8), 1834-1850.

67. Zang, B.; Lv, P.; Warren, C. M., Housing prices, rural–urban migrants' settlement decisions and their regional differences in China. *Habitat International* **2015,** 50, 149-159.

68. Long, Y.; Zhou, Y., Quantitative evaluation on street vibrancy and its impact factors: A case study of Chengdu. *New Archit* **2016,** 1, 52-57.

69. Li, C.; Wang, M.; Wang, J.; Wu, W. *The geography of city liveliness and land use configurations: Evidence from location-based big data in Beijing*; Spatial Economics Research Centre, LSE: 2016.

70. Wu, J.; Ta, N.; Song, Y.; Lin, J.; Chai, Y., Urban form breeds neighborhood vibrancy: A case study using a GPS-based activity survey in suburban Beijing. *Cities* **2018,** 74, 100-108.

71. Zheng, S.; Hu, X.; Wang, J.; Wang, R., Subways near the subway: Rail transit and neighborhood catering businesses in Beijing. *Transport Policy* **2016,** 51, 81-92.

72. Shen, Y.; Karimi, K., The economic value of streets: mix-scale spatio-functional interaction and housing price patterns. *Applied Geography* **2017,** 79, 187-202.

Chapter 2. Methods and data in the research of interactive relationship between urban development and population mobility

2.1 Introduction

There is an intricate relationship between urban development and population mobility. The population mobility between regions not only shows a certain dependence, but also has regional heterogeneity and endogeneity in the factors affecting population mobility. As the first law of geography put forward by Waldo Tobler, everything is related to other things, but things that are near are more related to things that are far away [1]. This rule requires more complex and in-depth models when studying population mobility and its influencing factors, because compared with the general statistical model, the model considering spatial correlation between cities makes the analysis results more accurate by incorporating spatial relations. Due to the different research scales, the statistical quality, statistical caliber and standard of city related data are quite different. The reliability, accuracy and timeliness of data are the basic issues to be considered in the model design. Therefore, for the national or regional scope, to explore the relationship between population mobility and urban development, we need to design the most appropriate model from the data sources, processing and analysis methods. In this context, combined with our research objectives and contents at different scales, this paper uses Geographic Information System (GIS), spatial econometric model and social network analysis to study the interaction between population mobility and urban development in Chinese cities from macro, meso and micro scales.

In recent years, with the development of information and communication technology, big data has become an important development direction and research field, and plays an important role in many disciplines. In the past, urban research in China has been restricted by data acquisition for a long time. This is because the data related to the city are highly dependent on the official surveying and mapping data, statistical data and official data of government administrative departments. However, since 2000, with the maturity of information and communication technologies (ICTs), especially the popularity of social media portals, smart phone applications and the birth of cloud computing, urban research has gradually changed from qualitative analysis to quantitative analysis. The so-called quantitative city research, refers to use in a variety of data on the certain theory basis and methods, to explore the general law of urban development, urban disease diagnosis problem, sim city operation and evaluating scientific research method of policy development and seek solutions, can be applied to support the status quo of urban planning and urban studies analysis, solution preparation and evaluation stages. The current quantitative urban research is not the application of planning informatization and new planning technology in the traditional sense, but pays more attention to the objective, intuitive and comprehensive analysis of urban phenomena, and transmits the analysis results to the government, planners, experts, scholars and urban residents through a variety of media, so as to improve the scientific of urban planning, urban management and related policy-making in many ways.

The emergence of big data provides a new basis for quantitative urban research. With the intervention of big data, many changes have taken place in the field of urban research. At the spatial scale, the data support makes the research change from small scale high progress and large range low precision to large-scale high-precision. On the time scale, data support makes urban analysis change from static section to dynamic continuous. From the perspective of local data, the focus of research has been shifted from “people-oriented” to “data-based”. In terms of research methods, data support makes urban analysis methods change from single team to open source crowdsourcing. Therefore, in this study, we combine traditional data with emerging big data to explore the interactive relationship between urban development and population mobility in different scales. The traditional data we used are a series of social and economic indicators which reflect the urban development status summarized by the National Bureau of Statistics at the macro level, which is considered to be the most complete and comprehensive official data in China. The big data we used include Tencent Location big data, Baidu heat map and the national dynamic survey dataset of population mobility released by the National Health Commission. After the relevant data is processed by privacy, different spatial econometric models are used to explore the interaction between urban development and population mobility.

2.2 Methods and applications of Geographic Information System (GIS)

2.2.1 Theories and development of GIS

GIS is a computer system for collecting, storing, analyzing, and displaying spatial information, and a general technique for processing and analyzing geographic data. In the integration of water quality models and GIS, it is used as a powerful tool for spatial discretization, parameterization and visualization of water quality models. Geographic information systems (GIS) combine a variety of technologies, processes and methods, are dependent on a variety of operations and a large number of applications, and are widely used in engineering, planning, management, transportation and telecommunications [2, 3]. GIS is a service system based on geographic location to analyze and visualize geographic data. Generally speaking, it is any data system that describes the integration, storage, editing, analysis, sharing, and display of geographic information. GIS systems allow users to create tools to interactively consult, analyze spatial information, edit data and maps, and display the results of all these operations. At present, geographic information science has been developed in many universities and become the basic subject of applied systems in urban studies and geosciences analysis [4, 5].

Geographic information system (GIS) was first proposed by Roger Tomlinson in 1968 [6]. In the early 1960s, Tomlinson also led the first computerized GIS on the Canadian Geographic Information System, known as the Canadian Geographic Information System, which was used to store, analyze, and process all the information collected on Canada's land inventory. The system is designed to simulate land use capabilities in rural Canada by mapping information about soil, agriculture, recreation, flora and fauna, and land use. The original Canadian Geographic Information System was the prototype for the development of modern geographic information systems. It provides coverage, measurement, digitization, and scanning through improvements to computer mapping applications. At the same time, the system supports geographic coordinates across the entire continent, encodes lines into truly embedded topological arcs, and stores important information about attributes and locations in separate files [7]. The system continued until about 1990 and established the first large digital land resource database in Canada. Subsequently, it was developed as a mainframe - based system to support the planning and management of federal and provincial resources. Its greatest advantage is the comprehensive and accurate data collection and analysis of complex ground information on a continental scale [5].

With the development of nearly half a century, geographic information system has developed into a powerful data software and applied in different disciplines [8]. The active GIS market and the continuous improvement of related software and hardware have promoted its wide application in science, government and industry, including real estate, urban planning, public health, defense and medical, natural resources, transportation and logistics, etc. Nowadays, with the development of electronic information technology, GIS is also extended to integrate spatial and temporal information of location services.

With the improvement of the system's functions and the expansion of its application scope, it has been more and more recommended as a tool for public participation. In the outline of the 12th Five-Year Plan for the development of surveying and mapping geographic information, the state has emphasized that the development of geographic information should be included in the national plan for strategic emerging industries. With the introduction of a number of guiding policies at the national level, geographic information system will usher in unprecedented development opportunities. Strong public location base, improve the collection, management, analysis, and the ability to output a variety of geospatial information, relying on strong complex analysis model, space comprehensive analysis ability and take the geographic research and geographic decision-making space of the human-machine interactive decision-making basis for the purpose of all make it in the analysis of urban space has a unique advantage.

2.2.2 Explanatory spatial data analysis (ESDA) in GIS

Although geographers have a long history of using spatial analysis methods to study geographical problems, the use of spatial analysis as an independent concept emerged with the development of GIS technology. At present, various spatial analysis techniques have been found, including spatial data manipulation, spatial data analysis, spatial statistical analysis and spatial modeling [9]. At present, spatial analysis presents a wealth of application fields, urban and regional fields, and forms a variety of connections with market, transportation and natural resources. Since the 1990s, with the popularization of personal computers, especially the maturity of GIS technology, a new platform has been provided for the spatial analysis of geographical phenomena and processes. Spatial data analysis methods based on GIS are becoming the focus of research. In the last decade of development, the key technologies of space analysis have changed greatly. Geographic information system and remote sensing technology have ensured the rich environment of space data. New analysis models and methods for dealing with space problems have been put forward constantly. Exploratory spatial data analysis (ESDA) method is a key step for statistical analysis of spatial data and mining the spatial structure and correlation of elements before modeling [10, 11]. So, in this chapter, we summarize the spatial data analysis method based on GIS platform which was used in the full paper.

Pearson correlation analysis

Correlation is a technique for investigating the relationship between two quantitative, continuous variables, for example, age and blood pressure. Pearson's correlation coefficient (r) is a measure of the strength of the association between the two variables [12]. The first step in studying the relationship between two continuous variables is to draw a scatter plot of the variables to check for linearity. The correlation coefficient should not be calculated if the relationship is not linear. For correlation only purposes, it does not really matter on which axis the variables are plotted. However, conventionally, the independent (or explanatory) variable is plotted on the x-axis (horizontally) and the dependent (or response) variable is plotted on the y-axis (vertically). The magnitude of the correlation is evaluated by the Pearson correlation coefficient, which is also refers to Pearson's r. The value of Pearson's r is between +1 to -1, where 1 refers to a total positive correlation, -1 refers to a total negative correlation, and 0 refers to no linear correlation. And it can be calculated through the following Eq. 2- 1.

$$r = \frac{n\sum xy - \sum x \sum y}{(n\sum y^2 - (\sum y)^2) * \sqrt{n\sum x^2 - (\sum x)^2}} \qquad \text{(Eq. 2-1)}$$

Where, x and y are values of variables, and n is size of the sample. The value of correlation coefficient can be interpreted in the following manner: If 'r' is equal to 1, then there is perfect positive correlation between two values; If 'r' is equal to -1, then there is perfect negative correlation between two values; If 'r' is equal to zero, then there is no correlation between the two values.

Multivariate ordinary least squares (OLS) regression

On the basis of correlation analysis between variables, the linear relationship between independent variables and dependent variables is found by using Multivariate ordinary least squares (OLS) regression [13]. Traditional OLS models assume that all independent variables are global and spatially fixed. Meanwhile, the relationships of OLS model can be expressed as follows:

$$y_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_j + \varepsilon_i \qquad \text{(Eq. 2-2)}$$

where β_0 represents the intercept, β_j represents the regression coefficient for the independent variable x_j, ε_i represents the error term.

In traditional OLS models, the determination coefficient is defined as the ratio of the explained variance to the total variance of the dependent variable, and is usually used for the degree of fitting between the predicted value and the actual value. R^2 is the square of the correlation coefficient, whose value ranges from 0 to 1. The closer to 1 is, the better the fitting degree of the model is. However, the empirical analysis shows that there are some errors in the correlation between variables only represented by R^2. This is because in the regression analysis, when variables are added continuously, R^2 is likely to remain unchanged, thus greatly affecting the selection of variables in the model. Therefore, we adopt the adjusted R^2 as the index of the explanatory ability between variables, which can capture the linear relationship between multiple variables in the model at the same time and reduce the errors in the empirical analysis. The adjusted R^2 can be calculated through the following equation:

$$Adjusted\ R^2 = 1 - \frac{n-1}{n-p} * (1 - R^2) \qquad \text{(Eq. 2-3)}$$

In addition, different criteria are also used to select different models, such as Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC) and Cross Validation (CV) [14-16]. When choosing different criteria, the optimal model will be set differently. We will illustrate the selection of different criteria with specific cases in the following chapters.

Spatial weight matrix

Spatial weight matrix is a key step of spatial data analysis and modeling. Spatial weight matrix integrates spatial relations directly into mathematical algorithms by spatial statistics. It measures the spatial relationship between research objects through a variety of expressions [17]. In exploratory spatial data analysis, there are a variety of weighted kernel functions to choose from, including inverse distance, fixed distance, K nearest neighbor, etc. As the core of the spatial weight matrix, the weighted kernel function can vary with the different research objects. In the software of GWR4 developed by Dr. Luc Anselin and his team [18], there have four type of kennels can be choose, including bi-square, adaptive Gaussian, fixed bi-square and fixed Gaussian, the detail of each kernels see Table 2-1.

Table 2-1. Four kennel type options available in the software of GWR4.

Kernel type	Equation
Fixed Gaussian	$w_{ij} = exp\left(\frac{-d_{ij}^2}{\theta^2}\right)$
Fixed bi-square	$w_{ij} = f(x) = \begin{cases} \left(1 - \frac{d_{ij}^2}{\theta^2}\right)^2, & d_{ij} < \theta \\ 0, & d_{ij} \geq \theta \end{cases}$
Adaptive bi-square	$w_{ij} = f(x) = \begin{cases} \left(1 - \frac{d_{ij}^2}{\theta_{i(k)}}\right)^2, & d_{ij} < \theta \\ 0, & d_{ij} \geq \theta \end{cases}$
Adaptive Gaussian	$w_{ij} = exp\left(\frac{-d_{ij}^2}{\theta_{i(k)}^2}\right)$

Note: i is the regression point index; j is the locational index.
w_{ij} is the weight value of observation at location j for estimating the coefficient at location i.
d_{ij} is the Euclidean distance between i and j.
θ is a fixed bandwidth size defined by a distance metric measure.
$\theta_{i(k)}$ is an adaptive bandwidth size defined at the kth nearest neighbor distance.

Among the four kinds of kernel functions, we need to choose the appropriate kernel function according to the different research objects. Among them, the adaptive bi-square kernel can reduce the bandwidth in the data-intensive place, expand the bandwidth in the data-scattered place, and show the clear range without the kernel weight being zero. Moreover, this kernel function has been widely used in urban studies. For example, Peng Wang et al. (2019) used a spatial econometric model with an adaptive bi-square kernel to study spatial-temporal variations in determinants of urban shrinkage in Japan [19]. Lee et al. (2019) explored spatial differences in determinants of obesity incidence in US cities [20]. Chen et al. (2019) investigated the development of spatial-temporal tourism in Jiangsu province, China [21]. In our macro analysis, 286 prefecture-level cities in China were taken as research objects (as shown in Fig. 2-1), which could be abstracted as 286 spatial information points. At the same time, spatial distance differences existed between cities (information points). Therefore, in this case, we choose the spatial weight matrix based on the adaptive bi-square kernel for spatial econometric analysis and modeling.

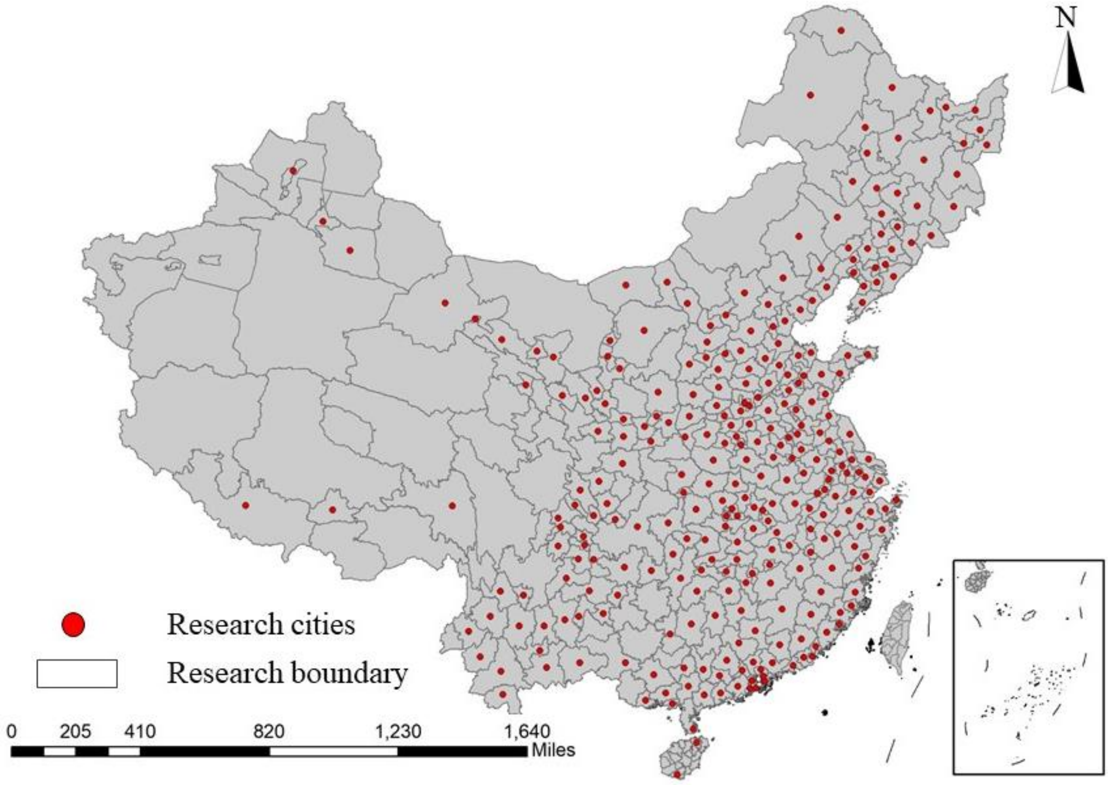


Fig. 2-1 The distribution of China's prefecture level cities in this paper.

In addition, the fixed Gaussian kernel has a fixed bandwidth, which is suitable for the research object of regular distribution. The spatial weight matrix based on fixed Gaussian kernel is widely used in the spatial analysis of fishing net data or remote sensing data [22, 23]. In our microscopic analysis, the main urban area of Qingdao is divided into a regular grid of 200 meters *200 meters to study the interaction between the urban built environment and population mobility, as shown in Figure 2-2. Therefore, in the analysis of chapter 6, we adopt the spatial econometric model based on fixed Gaussian kernel to carry out the research. Other kernel functions are not used in this paper due to the special model selection. The detailed model selection process is described in the following sections.

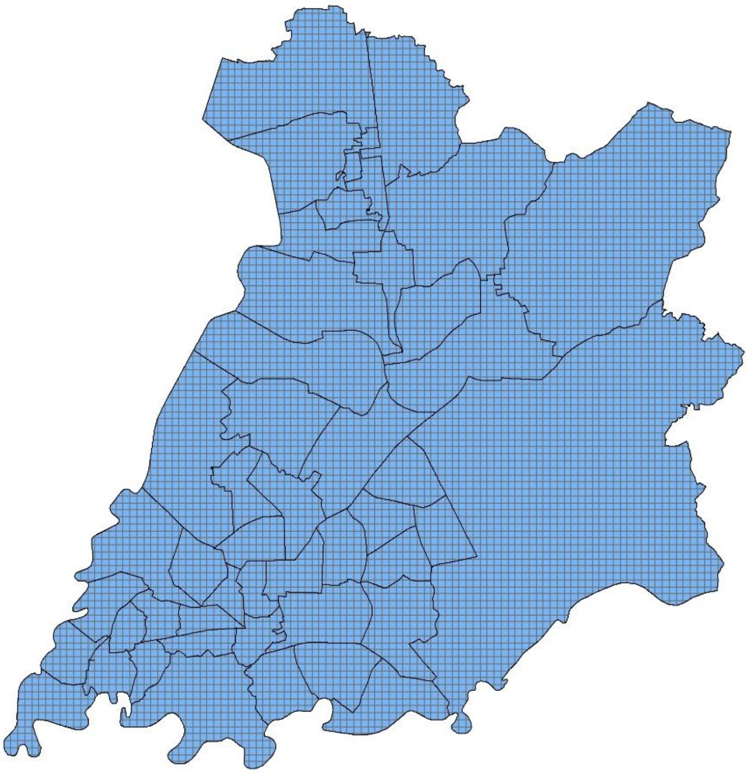
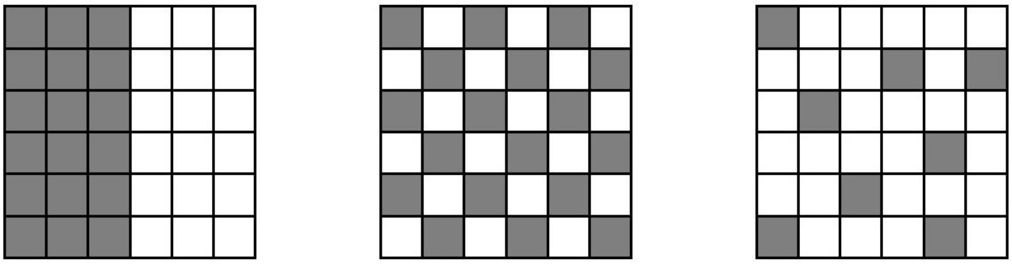


Fig. 2-2 200m fishnet in main urban area of Qingdao, China.

Spatial autocorrelation analysis

The traditional theory of spatial statistics is based on the assumption of independent observations. However, in the real world, especially when it comes to spatial data, independent observations are not universal in real life [24]. For data with geospatial attributes, it is generally believed that variables close to each other are more closely related than variables far away from each other in space [25]. The geographic and economic behaviors between regions generally have a certain degree of spatial correlation or spatial effect. However, in the actual analysis, many studies involving geospatial data ignore the effects of each other, and the wrong model setting leads to the deviation of the analysis results [26]. Spatial autocorrelation analysis is used to measure the basic attribute of geographic data: the degree of interdependence between data of one spatial location and data of other locations, which is considered as the most important attribute of spatial effect. Due to the influence of spatial interaction and diffusion, geographical data do not exist independently but interlinked. For example, the geographical and economic behaviors of labor force and capital flow influence and interact with each other in space, and the mutual influence of GDP development among all cities in the Yangtze River Delta region area, etc.

After taking spatial dependence into account, the pre-test of spatial autocorrelation must be carried out before the model is established for empirical analysis. If the spatial effect works, the spatial effect needs to be incorporated into the framework of the analytical model and estimated using appropriate methods. If there is no spatial effect, the model parameters can be estimated directly using general estimation methods (such as OLS). At the same time, after the introduction of spatial variables or the establishment of a spatial econometric model filtered by space, the quality of its effect also needs to be judged by the spatial correlation test, which can generally be realized by the spatial correlation test for the residual between the real value and the estimated value of the model. If the parameters tested do not show spatial correlation, then the model with spatial variables introduced or spatial effects considered has successfully dealt with spatial correlation. Fig. 2-3 shows three spatial autocorrelation types and samples. Generally speaking, there are three methods to examine the effects of spatial data, namely Moran's I [27], Geary's C and Getis index [28]. At present, the former is most widely used, so Moran's I is adopted in the spatial autocorrelation analysis in this paper. It should be pointed out that although there are differences in methods, the research contents can be divided into two categories: global spatial autocorrelation and local spatial autocorrelation.



Positive spatial autocorrelation | Negative spatial autocorrelation | No spatial autocorrelation

Fig. 2-3 Three spatial autocorrelation types and samples.

Global spatial autocorrelation is a rational number from -1 to 1 after normalized variance for spatial autocorrelation, so it is widely used to explain the effects between spatial data. Global Moran Index >0 represents positive spatial autocorrelation. The greater the value, the stronger the spatial correlation is. Global Moran index <0, the smaller the value is, the more obvious the spatial difference is; The global Moran index =0 indicates the random distribution of analysis elements without spatial effect. In this study, the absolute value of the value represents the degree of influence of the spatial effect (spatial aggregation or spatial dispersion). The global Moran index (GMI) can be calculated by the following formula:

$$GMI = \frac{n}{\sum_{i=1}^{n}\sum_{i\neq j}^{n} W_{ij}} * \frac{\sum_{i=1}^{n}\sum_{i\neq j}^{n} W_{ij}(x_i - x^*)(x_j - x^*)}{\sum_{i=1}^{n}(x_i - x^*)^2} \qquad \text{(Eq. 2-4)}$$

where x_i denotes the relative value of population mobility in city i, x^* denotes the mean value of x, n stands for the number of cities, and W_{ij} denotes the spatial weight matrix.

However, global spatial autocorrelation can only reflect the overall spatial effect, but cannot determine the specific location of elements with spatial effect [26]. Therefore, we also need to use the local Moran index to determine the local index of spatial correlation (LISA) [29]. The calculation formula of local Moran index (LMI) is as follows:

$$LMI = \frac{(x_i - x^*)}{\frac{1}{n}\sum_{i=1}^{n}(x_i - x^*)} \sum_{i=1}^{n}\sum_{i\neq j}^{n} w_{ij} \frac{1}{n}\sum_{i=1}^{n}(x_i - x^*) \qquad \text{(Eq. 2-5)}$$

After Local Moran's I classifies clusters/outliers, four results will be produced, namely high-high cluster (HH), high-low cluster (HL), low-high cluster (LH) and low-low cluster (LL). HH and LL reflect positive spatial correlation, while HL and LH reflect negative spatial correlation. In our study, HH indicates that the city shows a state of population inflow, and its neighboring cities also have similar characteristics. HL refers to the city where the population inflow, and population outflow in the surrounding cities. LH means the city where the population outflow, and the population inflow in surrounding cities. LL represents that both the city and its surrounding cities have population outflow.

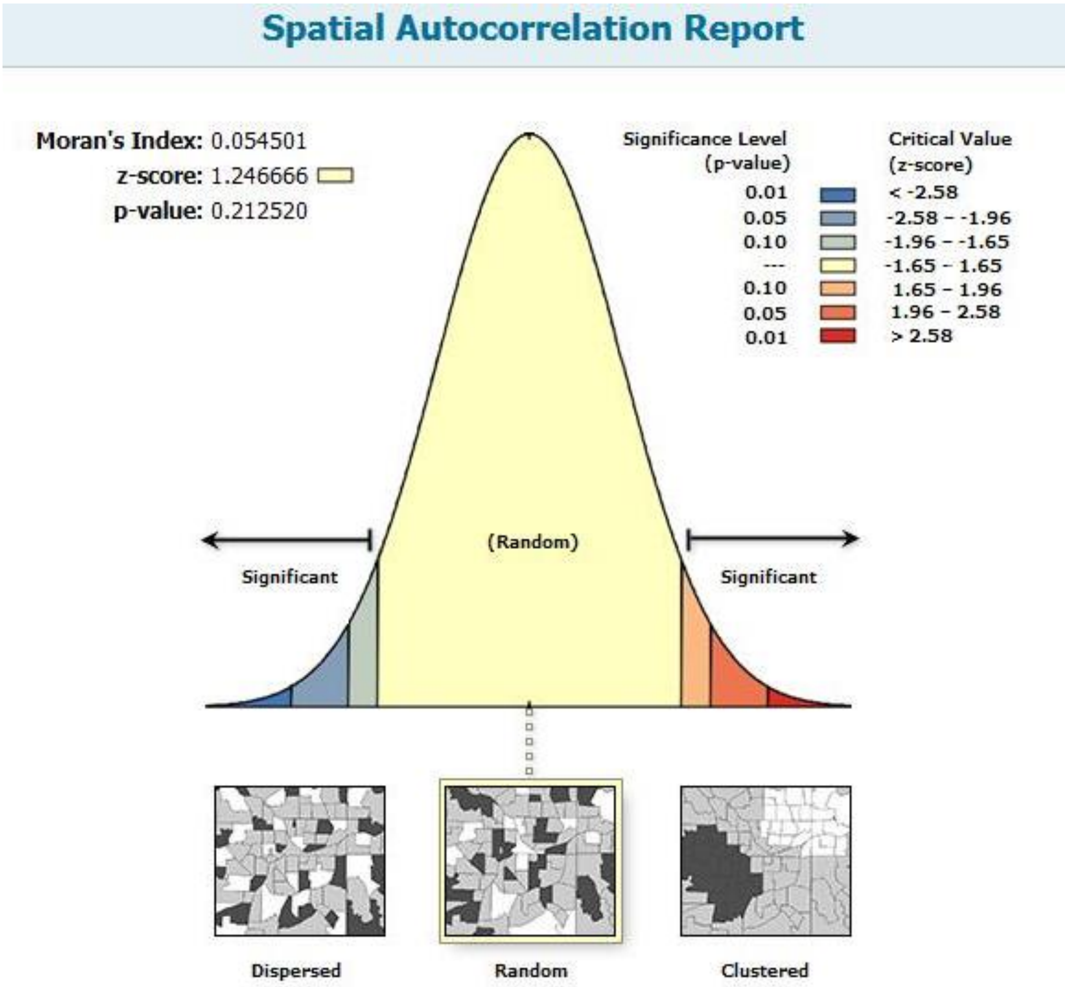


Fig. 2-4 Sample of autocorrelation analysis report in GIS.

The sample of autocorrelation analysis report is shown in Figure 2-4. The Z-score and p-value are two key parameters used to measure the Moran's I test significance level. The Z-score is calculated by the T-test.

2.3 Spatial econometric model

Tobler, an American geographer, put forward the first law of geography: everything is related, and the closer it is, the more relevant it becomes. This lays the foundation for spatial econometric analysis. The famous spatial econometrician Anselin pointed out that almost all data have spatial correlation and spatial dependence, and the development of spatial econometric model takes into account the internal structure of all spatial data. Goodchild believes that all spatial data has spatial correlation and spatial dependency [30]. The development of spatial econometric model is to add spatial structure to regional research, mainly to solve the problems of spatial dependence, spatial heterogeneity, spatial dynamics and spatial simulation. The appearance of spatial econometric model makes spatial analysis especially the field of spatial geographic economy analysis develop rapidly. As a complete branch, its application neighborhood includes real estate, geography, urban planning, economic geography and financial geography, etc. Not only this, the traditional fields of labor economy, energy economy, urban economy, environmental economy and industrial economy and other aspects of the research has also risen to the spatial level, the promotion of spatial econometric economy active in all aspects of social and economic research.

Anselin once defined spatial metrology as "a series of methods for studying various properties caused by space in the statistical analysis of regional scientific models". She explained regional science model is to point to in the model combines the related influence area, location and space, and are in estimates of the model and to determine the geographic reference data, the data may come from the point of the space, also can come from some areas, the former corresponds to the count coordinates, which corresponds to the relative position between the regions. In the model study, the difference of analysis results due to spatial effect should be considered. Compared with the traditional regression analysis, the definition of spatial econometric analysis is narrower. There are two types of spatial effects that can influence the analysis results in spatial econometric analysis: spatial dependence and spatial heterogeneity. Spatial dependence refers to the space between the main body behavior resulting from the interaction of a cross section of dependency, which means that the different location of the correlation between random variables or covariance structure mainly comes from the spatial organization form, the space organization form is by geography space relative position between the main body in the space (distance, space order). Spatial heterogeneity refers to the non-equilibrium of spatial structure, which is manifested as obvious spatial structural differences between subjects' behaviors. According to the different manifestation of spatial heterogeneity, spatial heterogeneity can be divided into spatial structure non-equilibrium and spatial heteroscedasticity. The non-equilibrium of spatial structure usually needs to set the spatial variable coefficient or spatial structure. Space heteroscedasticity usually requires heteroscedasticity treatment of the error term. Spatial heterogeneity can be divided into discrete heterogeneity and continuous heterogeneity [31]. Discrete heterogeneity displays spatial heterogeneity by setting regional dummy variables in the model. Continuous heterogeneity deals with spatial heterogeneity by setting the functional form of random spatial displacement variation of parameters.

Considering that our research objects and data have certain spatial attributes, we will adopt different spatial econometric models to study the interactive relationship between population mobility and urban development in this paper. Specifically, for different scales of research objects, we compare and apply the following spatial econometric models: spatial error model (SEM), spatial lag model (SLM), geographically weighted regression (GWR) model, semi-parametric geographically weighted regression (SGWR) model and multi-scale geographically weighted regression (MGWR) model. In the following sections, we will go into the details of each model.

Spatial error model (SEM)

The spatial error model (SEM) is a special case of regression with correlation of random interference terms, in which the non-diagonal elements of the covariance matrix represent spatial correlation structures [32]. The spatial structure can be specified in different ways and the error variance covariance matrix can be generated. The spatial autocorrelation between random interference terms may mean that there is a nonlinear relationship between independent variables and dependent variables [33, 34]. One or more regression independent variables are omitted in the regression model. The regression model should have an autoregressive structure. When the random interference item follows a process of spatial autoregression, that is, the random interference item at each position is a function of the random interference item at all other positions, then a form of spatial autoregression of the random interference item will be formed, and the spatial correlation will be introduced into this model, namely:

$$\varepsilon_i = \gamma \sum_j w_{ij}\varepsilon_i + \mu_i \qquad \text{(Eq. 2-4)}$$

Where, γ is the parameter of the space error, w_{ij} is the element in the ith row of the space weight matrix, and μ_i is assumed to be standard normal distribution. Formally, the spatial error model (SEM) can be expressed as:

$$Y = X\beta + \varepsilon, \varepsilon = \gamma w_\varepsilon + \mu \qquad \text{(Eq. 2-5)}$$

This model combines a standard regression model and a spatial regression model with random interference terms, and assumes that the error terms meet the conditions, that is, the variance is fixed and the error property is irrelevant. Since the mean value of the random interference term is 0, the mean value of the dependent variable is not affected by the spatial error no matter how the value of γ changes.

Because the spatial error model is similar to the sequential correlation problem in time series, it is also called the spatial autocorrelation model. The economic significance of the spatial error model lies in that the impact of an individual with another cross section will be transmitted to the adjacent individuals along with this special form of covariance structure. This form of transmission has a long-time continuity and attenuation, that is to say, the spatial influence has high order. In short, compared with the traditional regression model, the spatial error model considers the relationship structure of the error terms in space, which makes the analysis results of spatial data more reliable.

Spatial lag model (SLM)

The spatial lag model is mainly used to study the influence of the behavior of adjacent observed objects on the behavior of the whole system [35-37]. The spatial lag model includes explanatory variables X and spatial lag terms w_y, which can be formally expressed as:

$$Y = \rho w_y + X\beta + \varepsilon \qquad \text{(Eq. 2-6)}$$

in the formula, ρ is the spatial autoregression coefficient. If it passes the significance test, i.e., $\rho \neq 0$, it means that there is indeed an interactive relationship between regions. ε is the random interference term, which satisfies the condition that $E_{(\varepsilon)}=0$, $Cov(\varepsilon)=\alpha^2 * I$.

Because the spatial lag model is similar to the autoregressive model in the time series, the standard regression model integrates the spatial lag dependent variable, so the spatial lag model is also known as "mixed regression -- spatial regression model", which is referred to as the spatial autoregressive model for short. The economic implication of spatial lag model is that, if the variables concerned have spatial correlation expressed by spatial matrix, it is not sufficient to estimate and predict the variation trend of variables by only considering their own explanatory variables. Considering the influence caused by the spatial structure in the model, the influence caused by the spatial effect can be well controlled. To put it simply, compared with the traditional regression model, the spatial lag model considers the spatial effect of dependent variables, namely whether there is spatial autocorrelation, such as regional GDP in urban economics. When the spatial effect of dependent variables is included, the analysis results of spatial data can be more reliable.

Two points in particular should be noted at the same time. First, the use of spatial econometric model is required. It is possible to use spatial econometric model to deal with the problem of spatial effect only under the condition of the occurrence of spatial effect. If there are no obvious spatial problems, such as spatial dependence or spatial heterogeneity between research objects or research areas, more complex spatial econometric models need not be considered. The second is the choice of spatial econometric model. The choice of spatial error model and spatial lag model is not invariable. That is to say, for the problem of spatial effect, we should not be able to determine the model adopted at the beginning of the study, but have a responsible model screening process. The selection of the spatial econometric model starts with the most basic OLS model, and the Residual error of the OLS model is used to conduct the Lagrange multiplier test (LM). This test contains two statistics, namely, LM-error and LM-lag. If these two statistics are not significant, OLS will be selected as the final model. If only one statistic is significant, the LM-error statistic is significant to the spatial error model, while the LM-lag statistic is significant to the spatial lag model. If both of these statistics were significant, Anselin proposed a Robust Lagrange multiplier test, which also included two statistics, namely Robust LM-error and Robust LM-lag. If the two statistics are still significant at this time, the analysis model with the highest significance coefficient is selected. Fig. 2-5 describes in detail the selection process and criteria of the spatial error model and the spatial lag model.

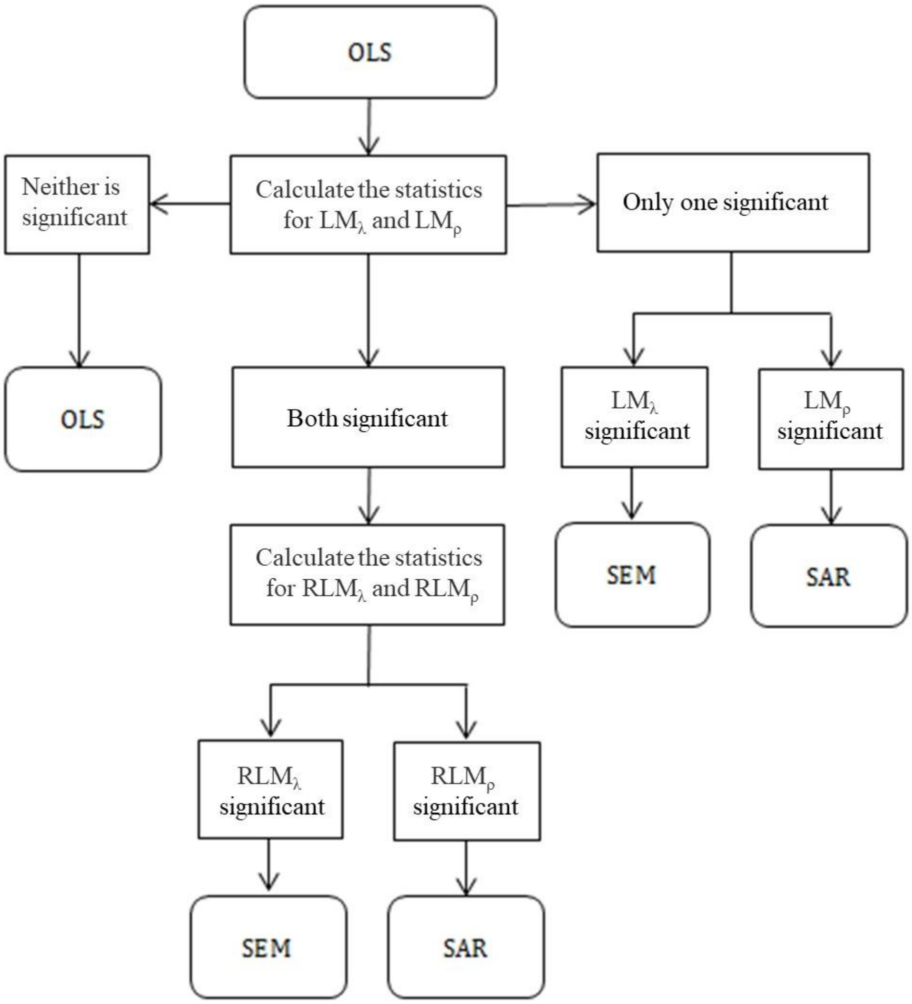


Fig. 2-5 The selection process and criteria of the spatial error model and the spatial lag model.

Geographically weighted regression (GWR) model

In city space analysis, observation data is usually in the N N group a different geographical location to obtain sample data, the global spatial regression model assumes that regression parameter has nothing to do with the sample data of geographic location, or in the space research area will remain stable and consistent, so N in different geographical location to obtain sample data, as in the same location for N sample data, the regression model is the same as the least squares regression model, the regression parameters is both optimal unbiased estimation of the observation point and global optimal unbiased estimation of observation point. In practical problems, we mentioned above that spatial data all have a certain spatial effect, namely, spatial heterogeneity. It means that regression parameters tend to behave differently in different geographical locations, that is, regression parameters change with geographical locations. In this case, if the global spatial regression model is still adopted, the estimated regression parameters will be the average value within the whole study area, which should not reflect the real spatial characteristics of regression parameters. In other words, the heterogeneity between spatial data is stable and captured, which affects the decision based on empirical analysis results. Based on this problem, foreign scholars proposed a regression model of spatial variable parameters, which incorporated the spatial structure of the data into the regression model and turned the regression parameters into a function of the geographical position of the observation point [38, 39]. Subsequently, geographically weighted regression (GWR) model has become a common method to solve this problem [40-43].

At present, the model has been widely used in meteorology, sociology, economics and other multidisciplinary research fields to effectively reveal the differences of spatial data. In the determined GWR model, spatial heterogeneity is taken into account the spatial location relationship of the data during the model fitting process. The local linear regression model for each observation point in the spatial dataset was calibrated using different weights [44]. The parameter in this model is a function representing the spatial position, which is obtained by weighting all adjacent observation values based on the function of distance attenuation [45].

The geographically weighted regression model is an extension of the ordinary linear regression model. The geographical location of the data from the observation point is embedded into the regression parameters. Its calculation formula is as follows:

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1} \beta_j(u_{i,} v_i) X_{ij} + \varepsilon_i \qquad \text{(Eq. 2-7)}$$

where (u_i, v_i) represents the coordinates of location i, $\beta_0(u_i, v_i)$ represents the intercept, $\beta_j(u_{i,} v_i)$ represents the local regression coefficient for the explanatory variable $y\ k$ at location i, and ε_i represents the error term. They are estimated using the local weighted least squares method. The selection of spatial weight matrix and kernel function is the core of the spatial econometric model, which has been described in detail in the previous part, and will not be explained here.

Semiparametric geographically weighted regression (SGWR) model

Compared with the traditional geographically weighted regression model, semiparametric geographically weighted regression [43, 46] pays more attention to the analysis of the diversity and nonstationarity of geospatial data. The model has been proven to be applicable in the fields of geography, environmental science, and economics [47, 48]. This model has better performance than the traditional geographically weighted regression model due to changes in the geographic parameters.

Generally speaking, not all regression parameters are the same or different in spatial position. The traditional OLS model assumes that all regression coefficients are invariable in spatial position, so the mean value is selected as the regression result of the study area, which has errors to some extent, as described in the previous part. At the same time, the GWR model assumes that the regression coefficients in the study area are non-stationary, i.e., there are differences between observation points. After considering the spatial heterogeneity, this method is more reliable than the traditional OLS model. However, the above two methods are considered the extremes of regression analysis, in complex urban spatial analysis, the variables and the difference are bigger, show the global impact of part variables and other variables show the influence of local differences, therefore, the above two methods are not able to fully describe the complex urban analysis content. Based on this, we adopt a semiparametric geographically weighted regression model to solve this problem. Compared with the former two methods, the latter can capture the change of regression coefficient at the global and local levels, and can deal with complex urban space problems. Therefore, the SGWR model is constructed based on the statistical yearbook data, and the formula is expressed by Equation (2-8):

$$y_i = \sum_{k=1} \beta_k x_k + \sum_{j=1} \beta_j \left(u_{i,} v_i\right) X_{ij} + \varepsilon_i \qquad \text{(Eq. 2-8)}$$

where k and j represent the global variables and local variables respectively. $\left(u_{i,} v_i\right)$ represents the coordinates of location i , $\beta_j\left(u_{i,} v_i\right)$ represents the local regression coefficient for the explanatory variable X_j at location i , and ε_i represents the error term.

Panel data regression model

There are two types of spatial data, one is sectional spatial data that only contains individual information, and the other is panel spatial data that includes time and individual information. The spatial econometric model mentioned above is mostly used to deal with cross-section data, but less applied to the analysis of panel spatial data. Compared with the cross-section data, the panel spatial data provides more data information after adding the time dimension, which increases the degree of freedom and reduces the collinearity problem between explanatory variables, so as to obtain more effective estimation results. Panel data can be used to address the problem of missing variables that do not change over time and represent individual heterogeneity. In panel data regression model, there are three kinds of models that can be selected and used, namely fixed effect regression model, random effect regression model and pooled regression model [49, 50].

The pooled model assumes that the coefficients of the intercept and explanatory variables are the same for all the individual members of the cross section, that is, it assumes that there is no individual influence or structural change on the members of the cross section [51]. Therefore, generally speaking, the pooled model is an OLS estimate developed after summarizing the data of the research object in all the study time. The characteristic of the mixed model is that the regression coefficient is the same regardless of any individual and interface. The formula of the pooled model is as follows:

$$y_{it} = \alpha + \beta_1 x_{1it} + \beta_2 x_{2it} + \cdots + \beta_k x_{kit} + \mu_{it} \qquad \text{(Eq. 2-9)}$$

Fixed effect model is an important analysis method in panel data regression. It refers to a variable method that only wants to compare the difference between independent variables or the effect of interaction terms with other independent variables. It emphasizes a class of variable methods that observe data changing with individuals but not with time. In general, the unobservable variables in fixed effects models are correlated with independent variables. The calculation formula of fixed effect model is as follows:

$$y_{it} = \alpha_1 + \beta_1 x_{1it} + \beta_2 x_{2it} + \cdots + \beta_k x_{kit} + \mu_{it} \qquad \text{(Eq. 2-10)}$$

where, intercept term α_1 means that there are different intercept terms for different cross section members, and its change is related to x_{kit}.

The coefficients of the random effects model are regarded as random variables and generally assume normal distribution. It has nothing to do with independent variables in the model. The calculation formula of random effect model is as follows:

$$y_{it} = \alpha_1 + \beta_1 x_{1it} + \beta_2 x_{2it} + \cdots + \beta_k x_{kit} + \mu_{it} \qquad \text{(Eq. 2-11)}$$

although the structural form of the random effects model and the fixed effects model is the same, the difference is that the random variable α_1 is independent of each explanatory variable x_{kit}.

It should be noted that both fixed effects model and random effects model are one of the methods of panel data regression. We cannot prejudge which model to use. Similar to the selection of spatial error model and spatial lag model, the selection of fixed effect model, random effect model and pooled model requires a screening process, although their emphasis may be different. In practice, there are several ways to determine the type of model used for panel data regression, such as Hausman test, which will be explained in more detail later.

Here are three tests to determine the optimal choice of the three models, they are F-test, BP -test and Hausman test respectively. Among them, F-test is used to compare and select fixed effect model and pooled model. If $p<0.05$, the null hypothesis cannot be rejected, that is, the fixed effect model is superior to the pooled model. The BP-test is used to judge the pooled model and the random effects model. If $p< 0.05$, the null hypothesis cannot be rejected, that is, the random effect model is superior to the pooled model. The Hausman test is used to select the fixed effect model and the random effect model. If $p<0.05$, the null hypothesis cannot be rejected, that is, the fixed effect model is superior to the random effect model. Combined with the three testing methods, we can select the optimal panel data regression model according to different research objects and data.

2.4 Social network analysis (SNA)

In this paper, in order to fully capture the spatial and temporal pattern of urban population mobility during the Spring Festival, we also adopted the social network analysis method. Social network analysis is a quantitative analysis method developed by sociologists based on mathematical methods and graph theory [52-54]. In recent years, this method has been widely used and played an important role in the fields of career mobility, the impact of urbanization on individuals, the world economic system and international trade. Social network analysis is a mature analytical method in the field of sociology, which can be used by sociologists to explain some sociological problems with ease. Experts in many disciplines, such as economics, management and other fields, begin to consider drawing on the research methods of other disciplines when facing many challenges in the new economic era -- the knowledge economy era. Social network analysis is one of them.

Network refers to all kinds of associations, and Social Network can be simply referred to as the structure constituted by Social relations [55]. In social network analysis problems originated from physics, adaptive network, through studying the relation network, help the relationships between individuals, "micro" network and large-scale social system "macro" structure, through mathematical method, graph theory and the quantitative analysis method, since the 1970 s in sociology, psychology, anthropology, gradually developed areas such as mathematics, communication science, a branch of study. As an important element of regional development, the movement track of population flow can reflect the correlation between cities based on "population flow" to some extent. Therefore, social network analysis is also used in the study of economic geography or urban geography and other related fields [56-60]. From the perspective of social network, the interaction of people in the social environment can be expressed as a pattern or rule based on relations, and the regular pattern based on such relations reflects the social structure, and the quantitative analysis of this structure is the starting point of social network analysis. Social network analysis is not only a tool, but also a relational way of thinking. It can be used to explain some problems in sociology, economics, management and other fields.

Network refers to all kinds of associations, and social network can simply be called the structure of social relations [61]. Therefore, from this aspect, social network represents a structural relationship, which can reflect the social relations between actors. The main elements that make up a social network are the following: actor, relation tie, dyad, triad, subgroup and group [62, 63]. An actor is not only a specific individual, but also a group, company, or other collective social unit. The position of each actor in the network is called “node”. The correlation between actors is called the relationship tie. There are various forms of relationships among people, such as kinship, cooperation, exchange, confrontation, etc., all of which constitute different ties. Dyad is the relationship consisting of two actors. This is a social network in its simplest or most basic form, and it’s the basis on which we analyze all kinds of ties. A triad is a relationship made up of three actors. Subgroups are subsets of any formal relationship between actors. A group is a collection of all actors whose relationships are measured. Social network analysis is a set of norms and methods to analyze the relationship structure and its attributes of social networks. It is also known as structural analysis, because it mainly analyzes the structure and attributes of social relations formed by different social units (individuals, groups or societies). In this sense, social network analysis is not only a set of techniques to analyze relationships or structures, but also a theoretical method—structural analysis. In the eyes of social network analysts, the object of sociology is the social structure, which is manifested as the relationship between actors. According to Barry Wellman, an analyst of social networks, "Network analysis explores deep structures -- patterns of networks that lie beneath the surface of complex social systems". For example, network analysts pay particular attention to how patterns of association in a particular network can influence people's actions by providing different opportunities or constraints [64, 65].

As a unique method for the study of social structure, B. Wellman summarized the methodological features of social network analysis in five aspects: (1) It explains people's behavior in terms of structural constraints on their actions, rather than through intrinsic factors (such as the socialization of norms), which see actors as pursuing desired goals in a voluntary and sometimes teleological manner. (2) It focuses on analyzing the relationships between different units rather than categorizing them according to their intrinsic properties (or nature). (3) It focuses on how the relationship forms composed of multidimensional factors jointly affect the behavior of network members, so it does not assume that there are only two-dimensional relationships among network members. (4) It views structures as networks between networks that may or may not belong to specific groups of people. It does not assume that groups with strict boundaries are necessarily barriers to structural formation. (5) Its analytical methods, which are directly concerned with the relational nature of certain social structures, aim to supplement - and sometimes even replace - mainstream statistical methods that require independent units of analysis. Therefore, according to the idea of social network analysis, any action of the actor is not isolated, but interrelated. The bond formed between them is the channel of information and resources transmission, and the network relationship structure also determines their action opportunities and results.

Social network analysis can be used to analyze social networks from many different perspectives, including centrality analysis, agglomerative subgroup analysis, core-periphery structure analysis and structural equivalence analysis, etc. Here we only introduce the first three methods used in this paper. Centrality is one of the focal points of social network analysis. The idea of what power, or centrality, an individual or organization has in its social network was one of the first topics discussed by social network analysts [66]. Centrality of an individual measures the degree to which an individual is in the center of a network, reflecting the importance of that point in the network. Therefore, as many actors/nodes in a network as there are individual centrality. In addition to calculating the centrality of individuals in a network, the central trend of the whole network can also be calculated. When some of the actors in a network are so closely related that they are combined into a subgroup, such a group is called a cohesive subgroup in social network analysis [67]. Condensing subgroup analysis is to analyze how many such subgroups exist in the network, the characteristics of the relations among the members within the subgroups, the characteristics of the relations between the members of one subgroup and another subgroup, etc. Due to the close relationship between the members of the condensed subgroup, some scholars also vividly called the condensed subgroup analysis "small group" analysis or "small world" analysis. The purpose of core-periphery structure analysis is to study which nodes in the social network are at the core and which nodes are at the edge. It has wide application and can be used to analyze the core and edge structure of elite network, scientific citation relationship network and organizational relationship network [68, 69].

Therefore, using the social network analysis method and taking the intensity of population flow between cities as the weight, we established a 290*290 directed weighting matrix P = (Pij) to characterize population mobility within 14 days. P_{ij} represents the intensity of population flow from city i to city j. We studied the network characteristics of population flow by using the social network analysis method. The population flow network is a small world, scale-free network between a fully regular network and a completely random network. Network characteristics are usually measured by the PageRank algorithm and "community (small world)" detection indicators.

$$P_{ij} = \begin{bmatrix} 0 & P_{12} & \dots P_{1(n-1)} & P_{1n} \\ P_{21} & 0 & \dots P_{2(n-1)} & P_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ P_{(n-1)1} & P_{(n-2)2} \cdots & 0 & P_{(n-1)n} \\ P_{n1} & P_{n2} & \cdots P_{n(n-1)} & 0 \end{bmatrix} \qquad \text{(Eq. 2-12)}$$

PageRank is an algorithm used by Google search engines to rank the importance of web pages. It is then applied to network analysis in any fields, such as bibliometrics, social network analysis and road networks [70, 71]. Compared with other centrality indices for evaluating nodes in network, such as degree, betweenness, and closeness, the PageRank algorithm not only consider the number of connections, but also measures the quality of connections, which means that if a node has fewer connections but all the important nodes are connected, it is still important. We believe that the mobility network formed during the Spring Festival is similar to the Internet, and cities with higher importance can attract more population and routes. Based on the intensity of population flow, we used the PageRank algorithm to rank the importance of urban nodes and get the hierarchical structure of population mobility. The formula of the PageRank algorithm is as follows:

$$PR_i = \frac{1-d}{N} + d\sum_{j\in B_j}\frac{PR_j}{L_j} \qquad \text{(Eq. 2-13)}$$

where PR_i is the PageRank value of city I; d is a constant, usually set to 0.85; N is the number of all cities; B_j is a collection of cities with all the population flow from city I; and L_j is the number of links from city i, which is weighted by the intensity of the population flow.

Many methods have been used for community detection testing, especially fast algorithms for large-scale networks, such as the Girvan-Newman algorithm, the CNM algorithm, SCAN algorithm, and so on [72-74]. In this paper, the multilevel algorithm was used for the community detection test, which is a bottom-up algorithm proposed by Blondel et al. through optimizing modularity [75].

2.5 Logistic regression model

Logistic regression, also known as logistic regression analysis, is a generalized linear regression analysis model commonly used in the fields of data mining, automatic disease diagnosis and economic prediction [76]. For example, discuss the risk factors of disease and predict the probability of disease occurrence based on the risk factors. Taking the analysis of gastric cancer as an example, two groups of people were selected, one was gastric cancer group and the other was non-gastric cancer group. The two groups of people must have different signs and lifestyles. Therefore, the dependent variable is gastric cancer, and the value is “yes” or “no”. The independent variable can include many things, such as age, gender, dietary habits, and helicobacter pylori infection. Independent variables can be either continuous or classified. Then, through logistic regression analysis, the weight of independent variables can be obtained, so as to roughly understand which factors are risk factors for gastric cancer. It also predicts a person's likelihood of developing cancer based on risk factors. The dependent variables of logistic regression can be binary or multi categorical, but the binary is more commonly used and easier to understand. Therefore, in the analysis of this paper, based on the dynamic monitoring data set of the national population mobility, we consider the factors affecting population mobility from the perspective of floating population.

If the linear regression model is directly buckled into the Logistic regression, the value interval of the two sides of the equation will be different and the general non-linear relation will be caused. Because the dependent variable in Logistic is dichotomous variable, the estimated value of a certain probability as the dependent variable of the equation ranges from 0 to 1. However, the value range on the right side of the equation is infinite or infinitesimal. That’s why Logistic regression was introduced. The essence of logistic regression is the logarithm of the probability of occurrence divided by the probability of no occurrence. It is this less complicated transformation that changes the contradiction of the value range and the curve relationship between independent variables of dependent variables. The reason is that the probability of occurrence and non-occurrence becomes the ratio, which is a buffer. The range of values will be expanded, and then the logarithmic transformation will change the whole dependent variable. Moreover, this transformation often makes the dependent variable and the independent variable present the linear relationship, which is summarized according to a lot of practice. Therefore, logistic regression fundamentally solves the problem of how to do if the dependent variable is not a continuous variable. In addition, the reason why logistic is widely used is that many practical problems are consistent with its model. For example, whether an event occurs in relation to other numerical independent variables.

In the logistic regression model, the dependent variable is set as Y, obeying binomial distribution, and the values are 0 and 1. Independent variables are X_1, X_2, X_3…….X_n. The formula of logistic regression model corresponding to independent variable is as follows:

$$P(Y=1)=\frac{EXP(\beta_0+\beta_1X_1+\beta_2X_2+\cdots\beta_nX_n)}{1+EXP(\beta_0+\beta_1X_1+\beta_2X_2+\cdots\beta_nX_n)} \qquad \text{(Eq. 2-14)}$$

As in the linear model, β_0 is a constant term and β_i is the partial regression coefficient corresponding to X_i (i=1,2,3,4,5…...n).

The ratio of the probability of an event occurring (P_i) and not occurring ($1-P_i$) is called the odds ratio, also known as the OR value [77, 78]. Because $0 < P_i < 1$, or is defined as a positive value, ranging from 0 to infinity. The linear model of logistic regression model can be obtained by logarithmic change of or value. The linear model of logistic regression model can be obtained by logarithmic change of OR value. The meaning of odds ratio is that the independent variable (X_n) changes by one unit, while the odds ratio corresponding to the dependent variable changes by EXP (β_n) units on average. Because the measurement units of different variables may be different, the absolute value of partial regression coefficient cannot be directly used to compare the relative effect of each independent variable, but standardized partial regression coefficient is needed. It is worth noting that the standardized regression coefficient is not used to construct the regression equation, it is only used to compare the contribution of different independent variables to the model, and the general regression coefficient β_0 is still needed to construct the regression equation model.

2.6 Data

2.6.1 Macro urban development statistics data

The macro urban development statistics data used in this paper are all from Urban Statistical Yearbook, including urban development yearbook, urban statistical yearbook, urban construction Yearbook and urban environment yearbook. These data are collected by national and local statistics bureaus, which are the most reliable source of urban development and population data in China. Urban Statistical Yearbook data include all aspects of macro urban development data, involving population, resources and environment, economic development, scientific and technological innovation, people’s life, public services and infrastructure. The specific macro urban development data used in each chapter will be introduced in each part.

2.6.1 Emerging big data

In this paper, we have used two kinds of emerging big data, Tencent location big data and Baidu heat map data, which will be respectively introduced in detail later.

Location-based services (LBS) obtain the geographical location of a mobile user through the wireless communication network or the external positioning method of network operators. When users allow various mobile applications to call LBS services, their movement trajectories will be accurately recorded in real-time through the positioning information. The movement of a single user in geographical space seems to be random, but may take on a specific pattern when a large population group is accessed. According to the statistical report on the development of the Internet in China, by the end of 2018, the number of instant messaging users had reached 792 million, and the number of mobile Internet users had exceeded 817 million, accounting for 98.6% of Internet users using mobile phones [79]. In this context, every smartphone user can be seen as a mobile sensor, reflecting social characteristics and allowing for the collection of a massive amount of individual movement information, in real time and in an efficient manner.

The dataset we used in our research is the “Migration Map” section of “Tencent Location Big Data” [65], with the time interval set to a day and the accuracy able to be traced back to the individual level (see Fig. 2-6). The website counts the number of changes in the location of the smart terminal within a certain time interval to filter, summarize, and count the data. In consideration of user privacy, the website only provides the total amount of population inflows and outflows in a day, with the city as the basic unit (the intensity of inflows, source, and outflows, limited to the destination of a single city on a certain day). The website provides a free Application Programming Interface (API) for researchers and programmers, allowing the above data to be obtained and used in scientific research. We used the API with Python programming language to obtain population mobility data during the Spring Festival of 2019 and store it in the SQL database.

The population mobility data we obtained contains the following content (after excluding the user's private information) and were added to a city as the basic units instead of the individual by Tencent company: the source city name and its coordinates, the target city name and its coordinates,

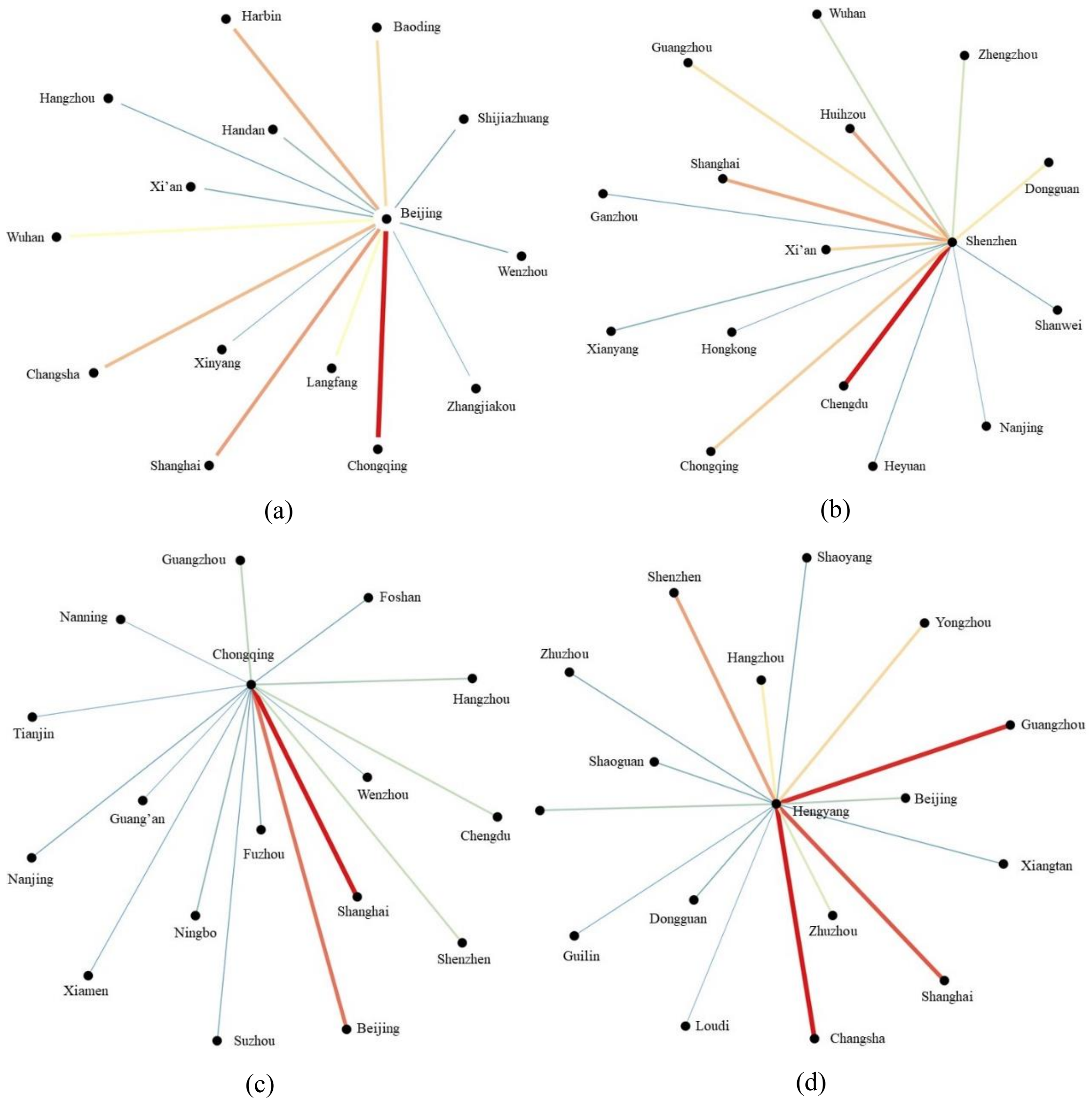


Fig. 2-6 Main population direction and intensity in four sample cities before and after the Spring Festival. (a, b, c, d) represent Beijing, Chongqing, Shenzhen and Hengyang, respectively. Beijing is the outflow direction before the Spring Festival, Chongqing is the inflow direction before the Spring Festival, Shenzhen is the inflow direction after the Spring Festival, and Hengyang is the outflow direction after the Spring Festival. The closer to red line color, the higher the intensity of population flow, and vice versa (original data).

time, mobility intensity, and mobility type, which is consistent with the content displayed on the website. After manual filtering and sorting, it contains a total of 40,591 pieces of information, each of which covers eight aspects (source city name and its coordinates, target city name and its coordinates, time, and mobility intensity). Based on this, we constructed a data table of 40,591 * 8, as shown in Table 2-2.

Table 2-2. Sample data of the Tencent location big data (after processed).

Sample data about outflows (a case in Beijing)							
Source	**Longitude**	**Latitude**	**Target**	**Longitude**	**Latitude**	**Date**	**Mobility intensity**
Beijing	116.4093	40.1841	Chongqing	107.8643	30.0553	2019/02/04	94193
Beijing	116.4093	40.1841	Changsha	113.1529	28.2295	2019/02/04	47416

Beijing	116.4093	40.1841	Baoding	115.1686	39.0222	2019/02/04	47404
Beijing	116.4093	40.1841	Langfang	116.5358	39.1117	2019/02/04	46473
Beijing	116.4093	40.1841	Harbin	127.9629	45.6369	2019/02/04	46416
…	…	…	…	…	…	…	…
Sample data about inflows (a case in Beijing)							
Source	**Longitude**	**Latitude**	**Target**	**Longitude**	**Latitude**	**Date**	**Mobility intensity**
Tangshan	118.3405	39.7358	Beijing	116.4093	40.1841	2019/02/04	10621
Shenyang	123.1366	42.0936	Beijing	116.4093	40.1841	2019/02/04	12807
Shanghai	121.4040	31.0844	Beijing	116.4093	40.1841	2019/02/04	69782
Hangzhou	119.4847	29.9049	Beijing	116.4093	40.1841	2019/02/04	28783
Nanchang	116.0244	28.6633	Beijing	116.4093	40.1841	2019/02/04	19356
…	…	…	…	…	…	…	…

Another emerging big data used in this paper is Baidu heat map data based on Baidu search engine. Baidu is the largest search engine and website in China. Its terminal market share accounted for 73.5%, covering 97.5% of all Chinese users, and the average daily response volume reached 6 billion times. In 2011, Baidu launched big data visualization products, namely Baidu heat map. Baidu heat map data is based on the location information obtained by mobile phone users when they access Baidu products (Baidu search, Baidu weather, Baidu map, and Baidu music, etc.), comprehensively calculate the heat value of population flow in different regions, and accurately reflect the degree of population aggregation in the region. Baidu heat map is updated every 15 minutes to obtain the dynamic information of population distribution in real time.

As a big data application with hundreds of millions of users, Baidu heat map has a wide range of coverage, high-resolution spatiotemporal characteristics and easy access, which embodies the potential and value in urban research. At present, the urban spatial analysis based on this data has been widely carried out, including urban spatial structure, urban job-housing balance, polycentric urban development, and urban population aggregation [80-83]. In this paper, the processed Baidu heat map data was used as the proxy of urban vitality.

Previous studies have shown that urban population has similar evolutionary patterns from Monday to Friday (weekdays), Saturday and Sunday (weekends). Therefore, this paper chooses December 14 (weekday) and 15 (weekend) in 2018 as the research time, calls the website application programming interface (API) through the Python programming language, obtains data at 1-hour intervals, and finally acquires a total of 36 Baidu heat maps (7:00-24:00). The population distribution data of spatial grid units generated after image processing is stored in the database.

2.6.1 National Migrant Population Dynamic Monitoring Survey data

National migrant population dynamic monitoring survey data is the annual large-scale national floating population sampling survey data of the National Health Commission since 2009, covering 31 provinces (autonomous regions, cities) and Xinjiang production and Construction Corps, where the floating population is relatively concentrated. The sample size is nearly 200000 every year, covering the basic information and floating range of floating population and family members. It includes the following aspects: surrounding and trend, employment and social security, income and expenditure and residence, basic public health services, management of marriage and family planning services, children mobility and education, psychological culture, etc. In addition, it also includes special investigation on social integration and mental health of floating population, special investigation on health and family planning service in outflow places, and special investigation on medical and health service for floating elderly.

The survey used stratified multi-stage random sampling method and probability proportional method to select samples from 348 cities covering 32 provincial units in China. In each city, the urban or suburban communities were randomly selected, and the floating population was investigated by face-to-face interviews in the selected communities. At present, this dataset has been widely used in the migrants' research [84-87]. In this paper, we start from the perspective of micro floating population to investigate the factors that affect the regional population flow. Therefore, the dataset gives us a good opportunity and a good entry point.

2.6 Summary

In this chapter, first, the concept and applications of GIS is introduced, including theory, development history and application method for ESDA. Second, the spatial econometric model is explained detail, including spatial error model, spatial lag model, geographically weighted regression model and semiparametric geographically weighted regression model. Then, we introduce the history and application of social network analysis method. Finally, it is the research data, we combine the macro data of urban development and the emerging big data, including Tencent location big data and Baidu heat map data, especially the national migrant population dynamic monitoring survey data, which gives us a good perspective to study the factors of population mobility. Please refer to each individual chapter for the specific application of methods and data.

Reference

1. Tobler, W. R., A computer movie simulating urban growth in the Detroit region. *Economic geography* **1970,** 46, (sup1), 234-240.
2. Chrisman, N. R., The role of quality information in the long-term functioning of a geographic information system. *Cartographica* **1984,** 21, (2), 79-87.
3. Tomlinson, R. F., *Thinking about GIS: geographic information system planning for managers*. ESRI, Inc.: 2007; Vol. 1.
4. Chang, K. T., Geographic information system. *International Encyclopedia of Geography: People, the Earth, Environment and Technology* **2016**, 1-10.
5. Cooperative, G.; Collins, F., The unique qualities of a geographic information system: a commentary. *Photogrammetric Engineering and Remote Sensing* **1988,** 54, (11), 1547-9.
6. Tomlinson, R. F., Geographic information systems—a new frontier. In *Introductory readings in geographic information systems*, CRC Press: 1990; pp 31-43.
7. Tomlinson, R. F.; HW, C.; DF, M., Computer handling of geographical data. **1976**.
8. Campagna, M., *GIS for sustainable development*. Crc Press: 2005.
9. O'sullivan, D.; Unwin, D., *Geographic information analysis*. John Wiley & Sons: 2014.
10. Dou, Y.; Luo, X.; Dong, L.; Wu, C.; Liang, H.; Ren, J., An empirical study on transit-oriented low-carbon urban land use planning: Exploratory Spatial Data Analysis (ESDA) on Shanghai, China. *Habitat International* **2016,** 53, 379-389.
11. Haining, R.; Wise, S.; Ma, J., Exploratory spatial data analysis. *Journal of the Royal Statistical Society: Series D (The Statistician)* **1998,** 47, (3), 457-469.
12. Zou, K. H.; Tuncali, K.; Silverman, S. G., Correlation and simple linear regression. *Radiology* **2003,** 227, (3), 617-628.
13. Hutcheson, G. D., Ordinary least-squares regression. *L. Moutinho and GD Hutcheson, The SAGE dictionary of quantitative management research* **2011**, 224-228.
14. Chen, S.; Gopalakrishnan, P. In *Speaker, environment and channel change detection and clustering via the bayesian information criterion*, Proc. DARPA broadcast news transcription and understanding workshop, 1998; Virginia, USA: 1998; pp 127-132.
15. Kohavi, R. In *A study of cross-validation and bootstrap for accuracy estimation and model selection*, Ijcai, 1995; Montreal, Canada: 1995; pp 1137-1145.
16. Sakamoto, Y.; Ishiguro, M.; Kitagawa, G., Akaike information criterion statistics. *Dordrecht, The Netherlands: D. Reidel* **1986,** 81.
17. Liu, Y.; Xiao, H.; Lv, Y.; Zhang, N., The effect of new-type urbanization on energy consumption in China: a spatial econometric analysis. *Journal of Cleaner Production* **2017,** 163, S299-S305.
18. Nakaya, T.; Charlton, M.; Yao, J.; Fotheringham, A. S., GWR4 User Manual: Windows application for geographically weighted regression modelling. *GWR4 Dev. Team* **2014,** 23, (14), 2019.
19. Peng, W.; Gao, W.; Yuan, X.; Wang, R.; Jiang, J., Spatiotemporal Differences in Determinants of City Shrinkage Based on Semiparametric Geographically Weighted Regression. *Sustainability* **2019,** 11, (24), 6891.
20. Lee, K. H.; Heo, J.; Jayaraman, R.; Dawson, S., Proximity to parks and natural areas as an environmental determinant to spatial disparities in obesity prevalence. *Applied geography* **2019,** 112, 102074.

21. Jin, C.; Xu, J.; Huang, Z., Spatiotemporal analysis of regional tourism development: A semiparametric Geographically Weighted Regression model approach. *Habitat International* **2019,** 87, 1-10.
22. Baraniuk, R. G.; Jones, D. L., Signal-dependent time-frequency analysis using a radially Gaussian kernel. *Signal processing* **1993,** 32, (3), 263-284.
23. Güngör, E.; Özmen, A., Distance and density based clustering algorithm using Gaussian kernel. *Expert Systems with Applications* **2017,** 69, 10-20.
24. Fischer, M. M.; Getis, A., Advances in spatial analysis. In *Recent developments in spatial analysis*, Springer: 1997; pp 1-12.
25. Anselin, L.; Getis, A., Spatial statistical analysis and geographic information systems. *The Annals of Regional Science* **1992,** 26, (1), 19-33.
26. Anselin, L., Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. *Geographical analysis* **1988,** 20, (1), 1-17.
27. Moran, P. A., Notes on continuous stochastic phenomena. *Biometrika* **1950,** 37, (1/2), 17-23.
28. Ord, J. K.; Getis, A., Local spatial autocorrelation statistics: distributional issues and an application. *Geographical analysis* **1995,** 27, (4), 286-306.
29. Anselin, L., Local indicators of spatial association—LISA. *Geographical analysis* **1995,** 27, (2), 93-115.
30. Goodchild, M. F., Geographical information science. *International journal of geographical information systems* **1992,** 6, (1), 31-45.
31. Anselin, L., Thirty years of spatial econometrics. *Papers in regional science* **2010,** 89, (1), 3-25.
32. Anselin, L.; Rey, S., Properties of tests for spatial dependence in linear regression models. *Geographical analysis* **1991,** 23, (2), 112-131.
33. Anselin, L.; Kelejian, H. H., Testing for spatial error autocorrelation in the presence of endogenous regressors. *International Regional Science Review* **1997,** 20, (1-2), 153-182.
34. Anselin, L.; Moreno, R., Properties of tests for spatial error components. *Regional Science and Urban Economics* **2003,** 33, (5), 595-618.
35. Anselin, L., Spatial econometrics. *A companion to theoretical econometrics* **2001,** 310330.
36. Anselin, L., Spatial externalities, spatial multipliers, and spatial econometrics. *International regional science review* **2003,** 26, (2), 153-166.
37. Anselin, L.; Bera, A. K.; Florax, R.; Yoon, M. J., Simple diagnostic tests for spatial dependence. *Regional science and urban economics* **1996,** 26, (1), 77-104.
38. Foster, S. A.; Gorr, W. L., An adaptive filter for estimating spatially-varying parameters: Application to modeling police hours spent in response to calls for service. *Management Science* **1986,** 32, (7), 878-889.
39. Gorr, W. L.; Olligschlaeger, A. M., Weighted spatial adaptive filtering: Monte Carlo studies and application to illicit drug market modeling. *Geographical Analysis* **1994,** 26, (1), 67-87.
40. Brunsdon, C.; Fotheringham, A. S.; Charlton, M. E., Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis* **1996,** 28, (4), 281-298.
41. Brunsdon, C.; Fotheringham, S.; Charlton, M., Geographically weighted regression as a statistical model. **2000**.
42. Fotheringham, A. S.; Brunsdon, C.; Charlton, M., *Quantitative geography: perspectives on spatial data analysis*. Sage: 2000.

43. Fotheringham, A. S.; Brunsdon, C.; Charlton, M., *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons: 2003.

44. Brunsdon, C.; Fotheringham, A.; Charlton, M., Geographically weighted summary statistics—a framework for localised exploratory data analysis. *Computers, Environment and Urban Systems* **2002,** 26, (6), 501-524.

45. Shi, H.; Laurent, E. J.; LeBouton, J.; Racevskis, L.; Hall, K. R.; Donovan, M.; Doepker, R. V.; Walters, M. B.; Lupi, F.; Liu, J., Local spatial modeling of white-tailed deer distribution. *Ecological Modelling* **2006,** 190, (1-2), 171-189.

46. Nakaya, T.; Fotheringham, A. S.; Brunsdon, C.; Charlton, M., Geographically weighted Poisson regression for disease association mapping. *Statistics in medicine* **2005,** 24, (17), 2695-2717.

47. Bitter, C.; Mulligan, G. F.; Dall'erba, S., Incorporating spatial variation in housing attribute prices: a comparison of geographically weighted regression and the spatial expansion method. *Journal of Geographical Systems* **2007,** 9, (1), 7-27.

48. Harris, P.; Brunsdon, C., Exploring spatial variation and spatial relationships in a freshwater acidification critical load data set for Great Britain using geographically weighted summary statistics. *Computers & Geosciences* **2010,** 36, (1), 54-70.

49. Hsiao, C., Panel data analysis—advantages and challenges. *Test* **2007,** 16, (1), 1-22.

50. Yaffee, R., A primer for panel data analysis. *Connect: Information Technology at NYU* **2003**, 1-11.

51. Plümper, T.; Troeger, V. E.; Manow, P., Panel data analysis in comparative politics: Linking method to theory. *European Journal of Political Research* **2005,** 44, (2), 327-354.

52. Carrington, P. J.; Scott, J.; Wasserman, S., *Models and methods in social network analysis*. Cambridge university press: 2005; Vol. 28.

53. Scott, J., Social network analysis. *Sociology* **1988,** 22, (1), 109-127.

54. Wasserman, S.; Faust, K., *Social network analysis: Methods and applications*. Cambridge university press: 1994; Vol. 8.

55. Knoke, D.; Yang, S., *Social network analysis*. Sage Publications: 2019; Vol. 154.

56. Giuffre, K., *Communities and networks: using social network analysis to rethink urban and community studies*. John Wiley & Sons: 2013.

57. Smith, D. A.; White, D. R., Structure and dynamics of the global economy: network analysis of international trade 1965–1980. *Social forces* **1992,** 70, (4), 857-893.

58. Song, J.; Feng, Q.; Wang, X.; Fu, H.; Jiang, W.; Chen, B., Spatial association and effect evaluation of CO2 emission in the Chengdu-Chongqing urban agglomeration: quantitative evidence from social network analysis. *Sustainability* **2019,** 11, (1), 1.

59. Ter Wal, A. L.; Boschma, R. A., Applying social network analysis in economic geography: framing some key analytic issues. *The Annals of Regional Science* **2009,** 43, (3), 739-756.

60. Zhong, C.; Arisona, S. M.; Huang, X.; Batty, M.; Schmitt, G., Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science* **2014,** 28, (11), 2178-2199.

61. Freeman, L., The development of social network analysis. *A Study in the Sociology of Science* **2004,** 1, 687.

62. Haythornthwaite, C., Social network analysis: An approach and technique for the study of information exchange. *Library & information science research* **1996,** 18, (4), 323-342.

63. Marin, A.; Wellman, B., Social network analysis: An introduction. *The SAGE handbook of*

social network analysis **2011,** 11.

64. Wellman, B.; Berkowitz, S. D., *Social structures: A network approach*. CUP Archive: 1988; Vol. 2.
65. Wellman, B.; Haase, A. Q.; Witte, J.; Hampton, K., Does the Internet increase, decrease, or supplement social capital? Social networks, participation, and community commitment. *American behavioral scientist* **2001,** 45, (3), 436-455.
66. Borgatti, S. P., Centrality and network flow. *Social networks* **2005,** 27, (1), 55-71.
67. Watts, D. J.; Strogatz, S. H., Collective dynamics of 'small-world'networks. *nature* **1998,** 393, (6684), 440-442.
68. Borgatti, S. P.; Everett, M. G., Models of core/periphery structures. *Social networks* **2000,** 21, (4), 375-395.
69. Rombach, M. P.; Porter, M. A.; Fowler, J. H.; Mucha, P. J., Core-periphery structure in networks. *SIAM Journal on Applied mathematics* **2014,** 74, (1), 167-190.
70. Jiang, B.; Zhao, S.; Yin, J., Self-organized natural roads for predicting traffic flow: a sensitivity study. *Journal of statistical mechanics: Theory and experiment* **2008,** 2008, (07), P07008.
71. Gupta, P.; Goel, A.; Lin, J.; Sharma, A.; Wang, D.; Zadeh, R. In *Wtf: The who to follow service at twitter*, Proceedings of the 22nd international conference on World Wide Web, 2013; 2013; pp 505-514.
72. Girvan, M.; Newman, M. E., Community structure in social and biological networks. *Proceedings of the national academy of sciences* **2002,** 99, (12), 7821-7826.
73. Clauset, A.; Newman, M. E.; Moore, C., Finding community structure in very large networks. *Physical review E* **2004,** 70, (6), 066111.
74. Xu, X.; Yuruk, N.; Feng, Z.; Schweiger, T. A. In *Scan: a structural clustering algorithm for networks*, Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, 2007; 2007; pp 824-833.
75. Blondel, V. D.; Guillaume, J.-L.; Lambiotte, R.; Lefebvre, E., Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* **2008,** 2008, (10), P10008.
76. Tranmer, M.; Elliot, M., Binary logistic regression. *Cathie Marsh for census and survey research, paper* **2008,** 20.
77. Harrell, F. E., Binary logistic regression. In *Regression modeling strategies*, Springer: 2015; pp 219-274.
78. Ozdemir, A., Using a binary logistic regression method and GIS for evaluating and mapping the groundwater spring potential in the Sultan Mountains (Aksehir, Turkey). *Journal of Hydrology* **2011,** 405, (1-2), 123-136.
79. China Internet Report 2019. https://www.abacusnews.com/china-internet-report (20 March),
80. Li, J.; Li, M.; Long, Y.; Dang, A., China polycentric cities based on baidu heatmap. *Shanghai Urban Planning Review* **2016,** 3, 30-36.
81. Tan, X.; Huang, D.; Zhao, X.; Yu, Y.; Leng, B.; Feng, L., Jobs housing balance based on Baidu thermodynamic diagram. *J. Beijing Normal Univ. Natural Sci.* **2016,** 52, (5), 622-627.
82. Wu, Z.; Ye, Z., Research on urban spatial structure based on Baidu heat map: A case study on the central city of Shanghai. *City Planning Review* **2016,** 40, (4), 33-40.
83. Lucang, W., Spatial-temporal characteristics of urban population aggregation based on baidu heat map in central areas of Wuhan city. *J. Hum. Settl. West China* **2018,** 33, 52-56.

84. Chen, M.; Wu, Y.; Liu, G.; Wang, X., City economic development, housing availability, and migrants' settlement intentions: Evidence from China. *Growth and Change* **2020,** 51, (3), 1239-1258.

85. You, Z.; Yang, H.; Fu, M., Settlement intention characteristics and determinants in floating populations in Chinese border cities. *Sustainable cities and society* **2018,** 39, 476-486.

86. Liu, Y.; Deng, W.; Song, X., Influence factor analysis of migrants' settlement intention: Considering the characteristic of city. *Applied Geography* **2018,** 96, 130-140.

87. Tan, S.; Li, Y.; Song, Y.; Luo, X.; Zhou, M.; Zhang, L.; Kuang, B., Influence factors on settlement intention for floating population in urban area: a China study. *Quality & Quantity* **2017,** 51, (1), 147-176.

Chapter 3. Spatiotemporal characteristics and driving forces of population mobility in urban China from 2000-2018

3.1 Introduction

Since the 1970s, the reform and opening up has accelerated the process of China's industrialization and urbanization. With the intensification of development differences between regions, the phenomenon of population mobility between urban-rural areas and between regions has become more and more serious [1]. For better employment, education, medical care and other social welfare, the rural population migrates to urban areas, and the population in less developed areas migrates to highly developed areas or metropolitan areas [2]. According to the 2018 report on the development of China's floating population issued by the National Health and Commission, by the end of 2018, the total number of interregional migrations in China was 241 million, accounting for about 17.2% of the country's total population. As the basis of all economic activities, population is essential for maintaining the sustainable development of the region. As one of the most widespread immigrants in the world, China's population mobility has played an important role in reshaping the national economic pattern [3, 4].

Population mobility is not only an important driving force to promote the process of urbanization and the growth of urban scale, but also the main force to promote regional economic development [5]. From 1978 to 2015, the contribution rate of population migration to China's economic growth was 19.99%, of which 57.79% was from 1978 to 1989. Based on the data of prefecture level cities in China, Wang [6] estimated that every 1% increase in the scale of population migration will increase the regional GDP by 0.54%. The population migration in Beijing, Shanghai and Guangzhou contributed more than 20% to the local economic growth, and the overall trend is rising. In 2015, the total amount of population migration accounted for about 35% of China's total labor force, about 15% higher than its share in China's total population [7]. At the same time, inter-regional migration is crucial to China's new urbanization and has a profound impact on promoting the development of urban-rural integration [8, 9].

In recent years, significant changes have taken place in China's population mobility. First of all, since 2010, the growth rate of population mobility scale showed a downward trend, and the mobility scale gradually reduced, indicating that China's population mobility scale began to enter the adjustment period after a long-term rapid growth [10]. Secondly, with the adjustment of industrial structure, labor-intensive industries and resource-intensive industries began to shift to the central and western regions, the process of labor force accompanying industrial flows will again optimize the allocation of human capital among regions [11, 12]. Furthermore, the central government has implemented strict migration control policies for large cities with an urban population of more than 3 million, which has caused significant changes in the scale of population mobility in first-tier cities [13-15]. Finally, as the scale of migration of the elderly population continues to increase and the scale of children migration declines, the rising living cost and the change of the inhabitant environment seem to have an increasing impact on population mobility [16-21].

At present, the phenomenon of population mobility in China has aroused extensive attention and discussion in academia [22-25]. However, there still have some limitations. First of all, China's urbanization rate exceeded 50% in 2010, becoming an important turning point in the demographic structure and urbanization process. At the same time, China's working-age population has peaked. However, few literatures have explored the issue of population migration after 2010. Second, there is a lack of comprehensive analysis of factors influencing population migration. The relevant literature only focuses on the analysis of the impact of economic factors or institutional factors on population mobility. Moreover, due to the spatial dependence and spatial heterogeneity of geographic data, traditional non-spatial regression methods may not only cause errors, but also fail to reveal the heterogeneity of driving forces in different spaces. Finally, the previous literatures do not reveal the changes of potential driving forces with time and space, which has important guiding significance for further urban development policy-making.

Based on the importance of population mobility and the shortcomings of current research, this chapter has three objectives: (1) Depict the spatial and temporal patterns of population mobility in prefecture level cities in China from 2000 to 2010 and 2010 to 2018. (2) Using spatial econometric model to explore the factors affecting population mobility from the perspective of urban development. (3) Assess the spatial and temporal changes of potential factors affecting population mobility. (4) Provide targeted suggestions for sustainable and coordinated development of cities and valuable insights for further research.

The reminder of this chapter is organized as follows. In 3.2, we systematically introduce the research framework which include study area, data and methods applied in the empirical analysis. 3.3 uses exploratory spatial data analysis method to study the changes in the spatial and temporal pattern of population migration from 2000 to 2010 and 2010 to 2018, and explores the potential driving forces and their changes with semiparametric geographically weighted regression model (SGWR). Finally, in 3.4, we conclude with an in-depth discussion of the main findings of the study and proposes feasible policies and development schemes as the summary.

3.2 Research framework

3.2.1 Research area and data

The sample we studied includes the current 290 prefecture-level and above cities in China, including 4 municipalities directly under the central government, 35 provincial capital cities and 251 general prefecture-level cities. Due to the lack of data, Taiwan, Hong Kong, Macau and some ethnic autonomous prefectures in western China are not included in this research. Fig. 3-1 shows the research area in this chapter.

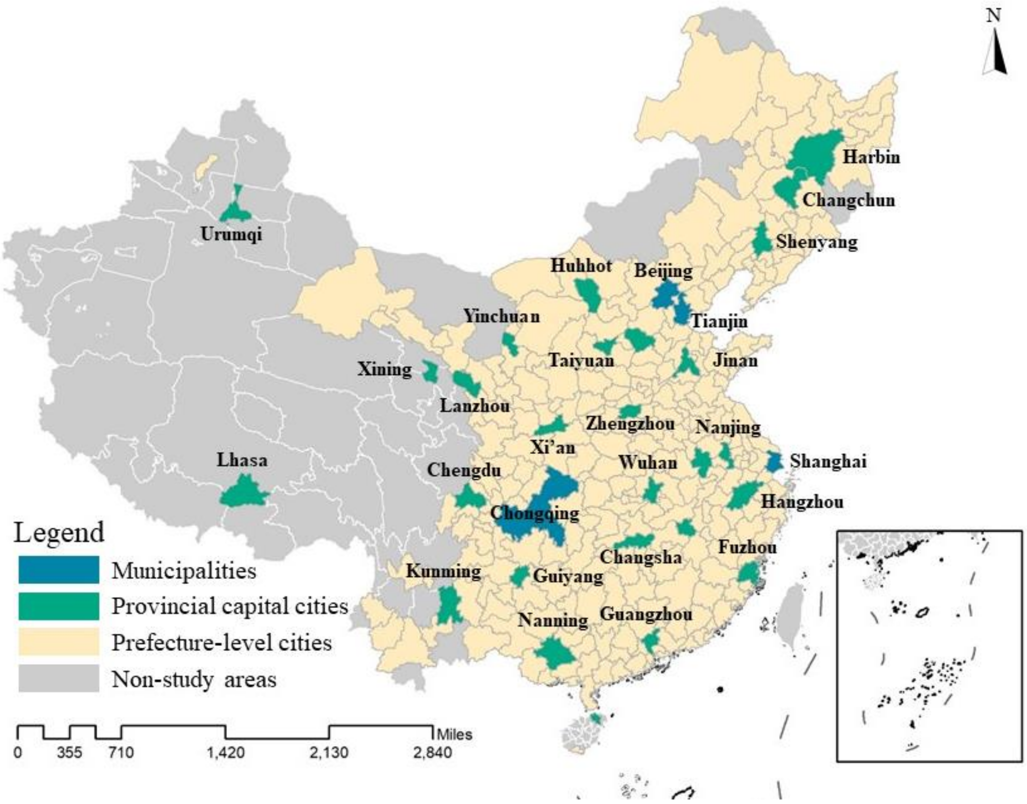


Fig. 3-1. Research areas in this study.

The population data used in the study are from the fifth and sixth population census data, the China population and employment statistical yearbook. The city type data are from China's city statistical yearbook, China's environmental statistical yearbook and China's regional economic statistical yearbook, some supplementary data are from the local statistical yearbook and the statistical bulletin of national economic and social development.

3.2.2 Definitions

Previous studies have not given a clear definition of population mobility in China, which may cause confusion. At present, scholars mostly conduct research from two aspects: rural-urban mobility and inter-regional mobility. In this chapter, we only focus on inter-regional population mobility. However, in China, the phenomenon of population mobility also exists in different districts of the same city, which cannot be understood as the inter-regional mobility. Therefore, population mobility should have a spatial dimension standard, that is, the population living in different registered cities and living in the city for at least 6 months. Statistically speaking, the migrant population in a city comes from other cities in the same province or cities in other provinces, which is similar to the floating population defined by the National Bureau of Statistics (NBS). It is worth noting that the definition of population mobility in this chapter is slightly different from that in the next chapter, because the corresponding study time in these two chapters is different. Therefore, in this chapter, we use population migration to represent the spatiotemporal behavior of the population.

3.2.3 Methods

In this study, we first used exploratory spatial data analysis (ESDA) methods to depict the changes in the spatiotemporal patterns of population mobility and verify the spatial dependence of population mobility. Then the ordinary least squares (OLS) and correlation test were employed to determine the driving force related to population mobility. Three types of regression analysis were further compared, including ordinary least squares (OLS), geographically weighted regression (GWR) and semiparametric geographically weighted regression (SGWR). Finally, the semiparametric geographically weighted regression (SGWR) was selected to reveal the differences in the driving force of population mobility in spatial and temporal dimension. Fig. 3-2 gives a flowchart of this research.

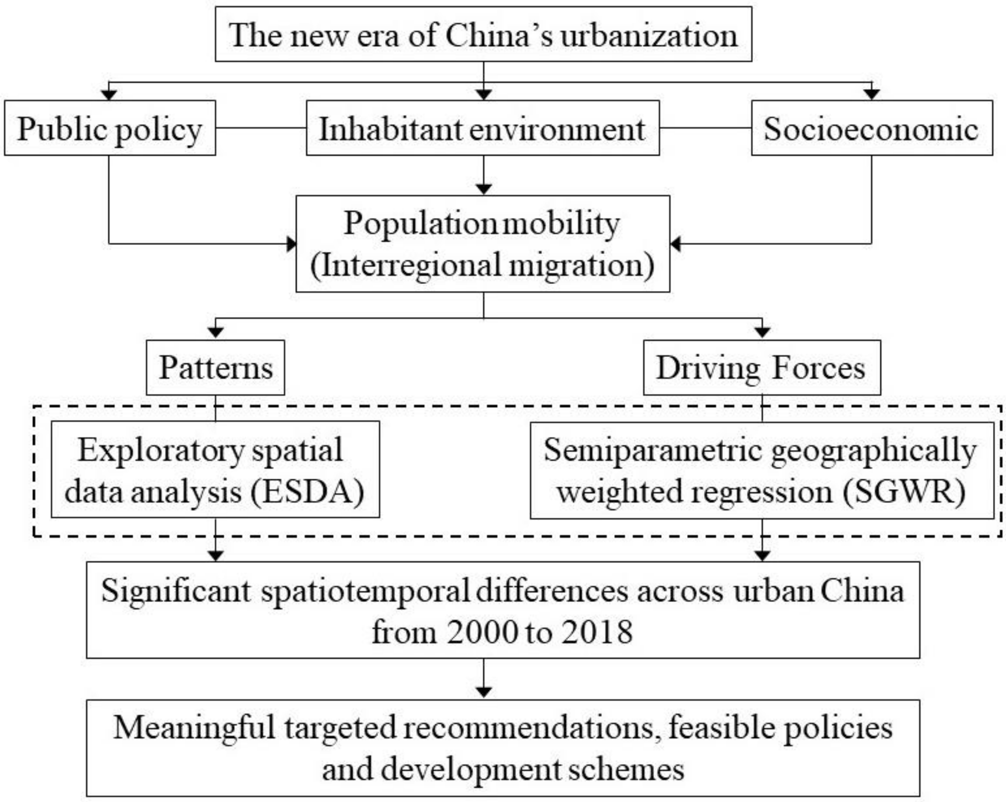


Fig. 3-2 Flow chart of the research.

3.2.4 Independent Variables Selection

Table 3-1 lists the detailed information of the explanatory variables used in this chapter. Previous studies believe that wage levels, as a representative of regional economic development, are an important force in promoting population mobility [26, 27]. With the transformation of China's economy, the service industry is considered the largest employment sector in the city. At the same time, foreign-funded enterprises are also showing great advantages in attracting population. Therefore, we selected four indicators of wage level, urbanization stage, globalization and industrial upgrading as the representative of urban economic development.

According to the literature review, population mobility laws and policies play an important role in controlling population migration. We introduced the dummy variable of migration control policy as a proxy for the national system. Meanwhile, local government expenditure and intervention in market management indirectly affect economic development, which in turn affects population migration. Generally speaking, cities facing huge financial pressures are often not very attractive to the population. Therefore, we selected four indicators of migration control policy, local government expenditure, government intervention index, and government financial pressure as representatives of urban public policies.

We have mentioned in the literature review that changes in inhabitant environment can affect population mobility. After considering the level of public service and living conditions, we take nine indicators, including traffic, road, urban green space, education, health, life, infrastructure investment, house price and air pollution as the representative of inhabitant environment.

Table 3-1. The detailed information of the explanatory variables in this research.

Categories	Variables	Definitions and description	Notation	Units
Public policy	Migration control policies	Dummy variable. 1 is the city that has implemented strict immigration policy, and 0 is used for other cities.	MICP	/
	Local government expenditure	Total expenditure at the end of the year.	LOGE	million yuan
	Government intervention index	The proportion of the difference between fiscal revenue and fiscal expenditure in fiscal expenditure. This indicator assesses the strength of local government intervention in the economy.	GOII	%
	Government financial pressure	Total fiscal deficit at the end of the year.	GOFP	million yuan
Socio-economic	Marketization	The Proportion of employees in non-state-owned enterprises in total employment.	MARK	%
	Wage level	Average wage of employees.	WAGE	10,000 yuan
	Urbanization stage	The Proportion of service industry employees in total employment.	URBA	%
	Globalization	The proportion of employees of foreign-funded enterprises in total employed employees.	GLOB	%

	Industrial upgrading	Proportion of GDP in secondary and tertiary industries.	INUP	%
	Transport	The number of public transport vehicles per 10000 residents.	TRAN	/
	Road	Per capita urban road area.	ROAD	m^2
Inhabitant environment	Green park space	Per capita green park space area.	GRLS	m^2
	Education	The number of full-time teachers in primary and secondary schools per 10000 residents.	EDUC	/
	Life	The number of recreational and sports facilities per 10000 residents.	LIFE	/
	Health	The number of professional doctors per 10000 residents.	HEAL	/
	Infrastructure	Investment amount of municipal public infrastructure.	INFR	million yuan
	House price	Average urban house price at the end of the year.	AUHP	10,000 yuan
	Air pollution	Annual average concentration of inhalable fine particles.	AIRP	Microgram / m^3

In order to improve the accuracy of the MGWR model, we performed multicollinearity test and correlation analysis on the selected explanatory variables. First, we perform OLS regression to detect multicollinearity between variables. After normalizing all the variables that conform to the normal distribution, the variance inflation factor (VIF) of each variable was calculated, and the variables with VIF > 7.5 were eliminated. In this process, the VIF of all variables is less than 7.5, indicating that there is no multicollinearity. Second, we conduct correlation analysis on variables and exclude variables that are not related to population migration at a confidence 95% confidence. In this process, some variables are excluded. Therefore, the remaining variables were used for SGWR model. Table 3-2 shows the results of VIF and correlation test.

Table 3-2. The results of VIF and correlation test of the variables.

	2000-2010			2010-2018		
	VIF	Coefficient	Sig.	VIF	Coefficient	Sig.
MARK	1.341	0.671	0.308	1.126	0.015	0.336
WAGE	3.424	0.602	0.000	2.695	0.540	0.000
URBA	1.745	-0.188	0.000	1.402	-0.088	0.508
GLOB	3.348	0.408	0.000	1.978	0.218	0.000

INUP	2.227	0.470	0.000	1.061	0.021	0.304
MICP	-	-	-	3.063	0.405	0.000
LOGE	4.051	0.583	0.000	4.921	-0.564	0.000
GOII	2.787	-0.163	0.000	1.611	0.260	0.000
GOFP	2.332	-0.235	0.000	2.489	-0.731	0.000
TRAN	1.693	0.527	0.000	2.613	0.643	0.000
ROAD	1.049	-0.132	0.306	3.417	0.503	0.000
GRLS	1.280	0.304	0.000	3.287	0.580	0.000
EDUC	1.373	0.346	0.000	1.512	0.357	0.000
LIFE	1.233	0.321	0.000	1.323	0.157	0.000
HEAL	2.128	0.664	0.000	4.372	0.739	0.000
INFR	2.161	0.389	0.000	1.799	0.391	0.000
AUHP	4.099	0.638	0.000	3.455	0.581	0.000
AIRP	1.052	-0.013	0.450	1.171	-0.130	0.000

3.2.5 SGWR model construction

For the SGWR model, the spatial weight matrix is very important. The choice of the spatial weight function has a greater impact on the parameter estimation of the model [28]. In this chapter, based on the following two points, we use an adaptive bi-square kernel instead of a fixed kernel to calculate the weight matrix: (1) Each city as a regression point is randomly distributed in the study area, and the adaptive kernel makes the dataset large enough for each local regression [29]. (2) The adaptive kernel can reduce the bandwidth of data-intensive areas and expand the bandwidth of data scattered locations, which has been widely recognized in the research [30, 31]. Meanwhile, the bandwidth of the weight function can affect the accuracy of the model to a large extent. The Akaike Information Criterion (AICc) and cross-validation (CV) are two common methods for determining bandwidth. Compared with the latter, the former can quickly and effectively resolve the difference in degrees of freedom in the model [32]. Therefore, the AICc was selected to determine the appropriate bandwidth when constructing the SGWR model. In this chapter, the significance ($p<0.05$) of all variables was defined as the pseudo t (Est/SE) > 1.96 or < -1.96 [32].

3. 3 Empirical results

3.3.1 Spatiotemporal patterns of population migration

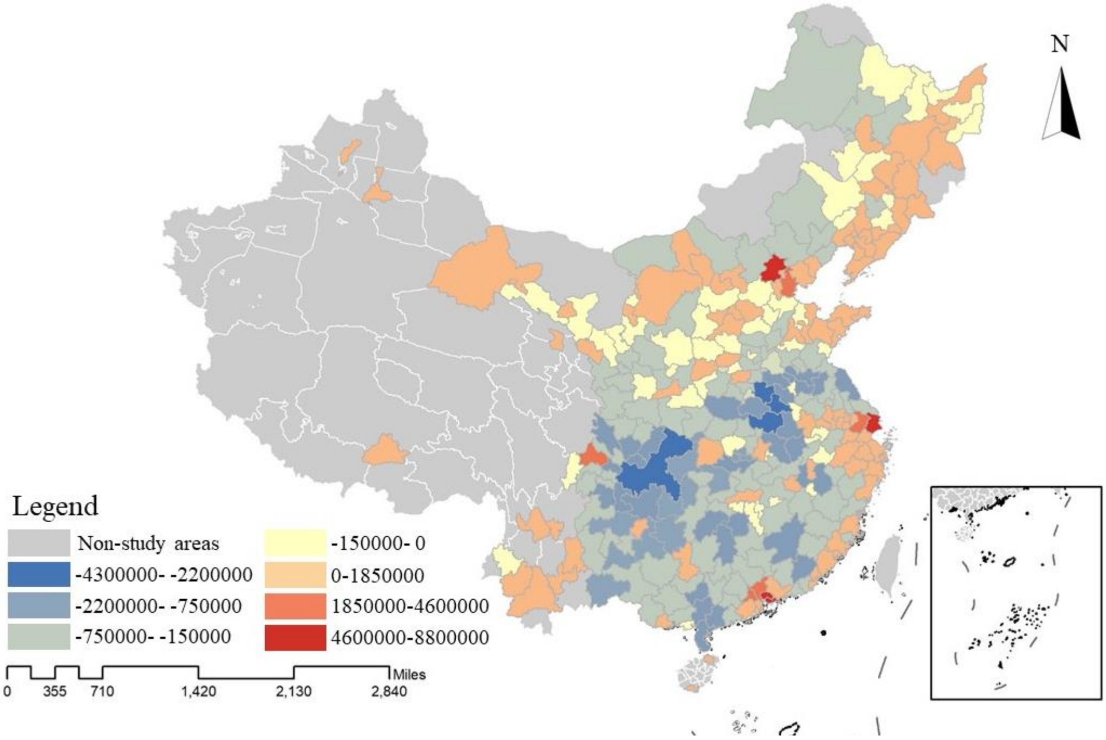


Fig. 3-3 The scale of net population migration of prefecture-level cities in China in 2000 (positive values represent population immigration; negative values represent population emigration).

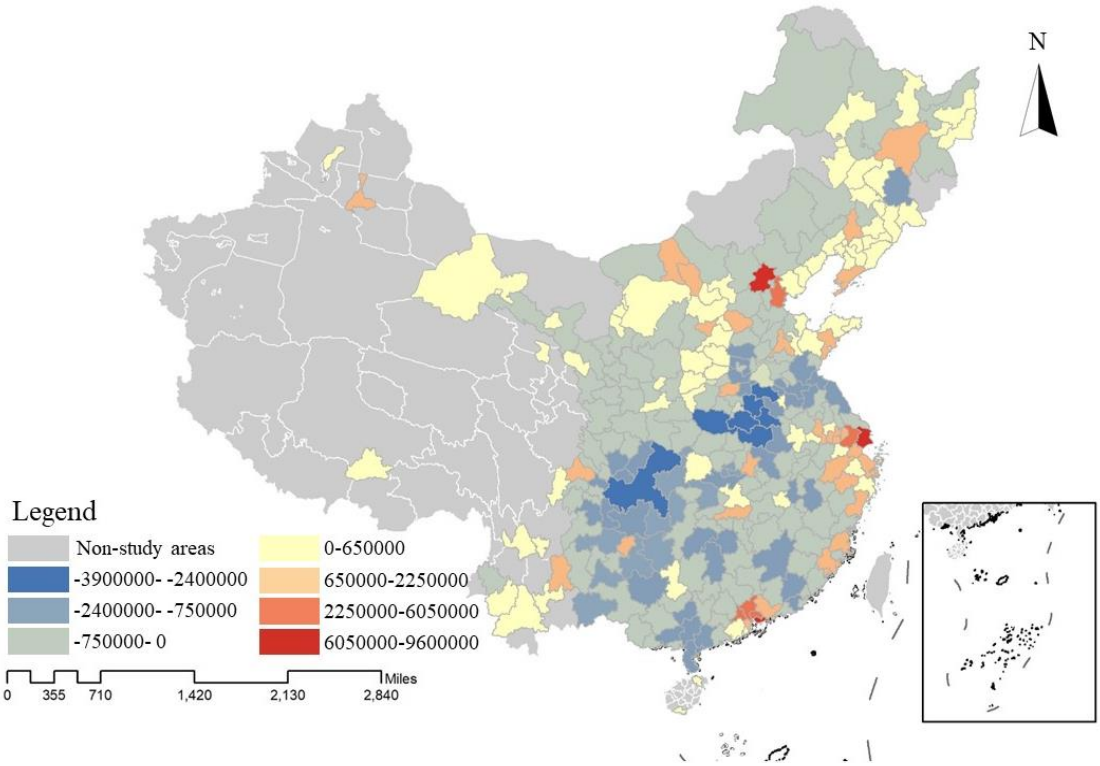


Fig. 3-4 The scale of net population migration of prefecture-level cities in China in 2010 (positive values represent population immigration; negative values represent population emigration).

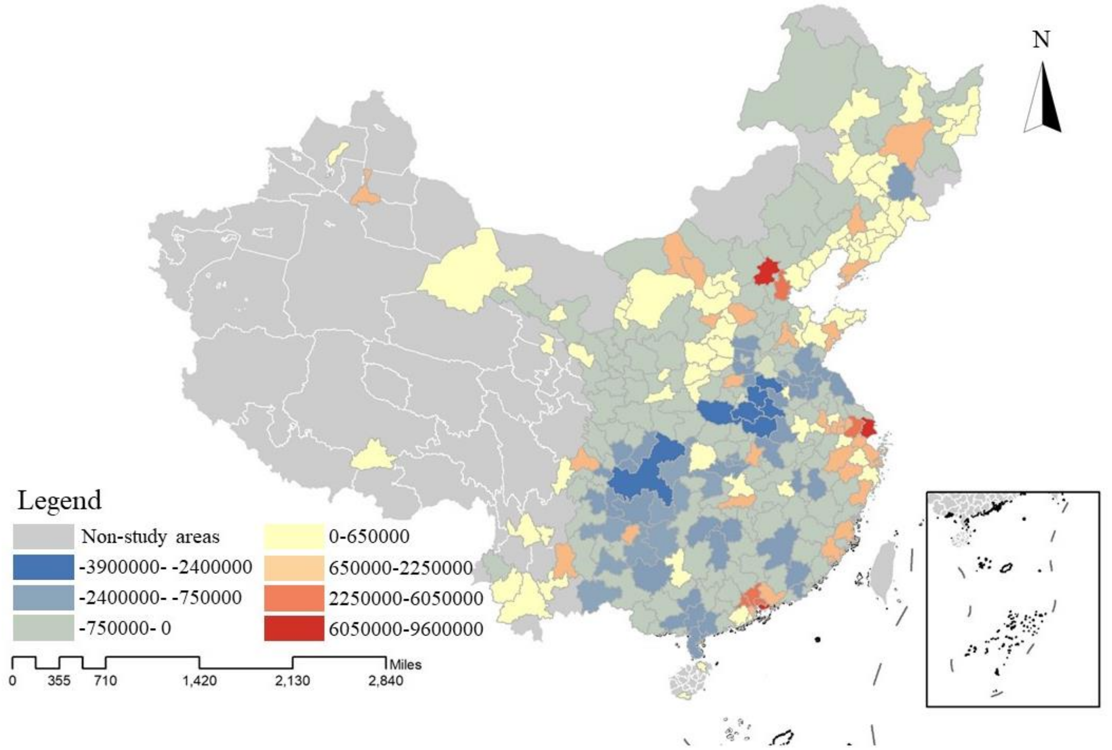


Fig. 3-5 The scale of net population migration of prefecture-level cities in China in 2018 (positive values represent population immigration; negative values represent population emigration).

Figs. 3-3, 3-4 and 3-5 are the net migration population size of prefecture-level cities in China in 2000, 2010, and 2018, respectively. Positive values represent population immigration, negative values represent population emigration. In 2000, the destinations of population migration were concentrated in the central cities of the Pearl River Delta (PRD), the Yangtze River Delta (YRD), and the Beijing-Tianjin-Hebei (BTH) region. Among them, the migration population of Shenzhen, Dongguan, Shanghai and Guangzhou exceeded 3 million, and that of other eastern coastal cities also exceeded 500000. After the reform and opening up, the eastern coastal cities gave full play to their location advantages, and under the support of the central government's policies, they have made considerable progress in economic development and become the preferred destination for population migration. At the same time, the central provinces have shown obvious population outflows. The emigration population scale of some cities in Henan and Guangxi reached 1 million, becoming the most serious areas of population loss. With the support of national strategies such as “Support Xinjiang and Tibet”, “Western Development”, and “Northeast revitalization”, the immigrant population of the above-mentioned areas increased slightly. In 2010, there were some changes in the scale of net migration population between cities. First, the scale of population migration in Northeast China, Xinjiang, and Tibet is steadily increasing. Secondly, Chongqing has become the city with the most serious population outflows, and the scale of the emigration population has exceeded 4 million. Finally, the scale of population migration in coastal open cities has increased significantly, and the surrounding cities of regional central cities have become new migration destinations, such as Jiaxing, Wenzhou, Ningbo and Qingdao, the scale of population migration has exceeded 1 million.

Figs. 3-6 and 3-7 show the statistics on the change and average annual growth rate of the population migration scale of prefecture-level cities from 2000 to 2010. Between 2000 and 2010, the migration scale of the eastern coastal cities exceeded 1 million. Shanghai has the largest migration population, reached 5.7 million. In terms of the average annual growth rate, the average annual growth rate of the migrant population in the coastal open cities exceeded 10%, which shows that the eastern region has become the destination of migration during this period. Although the average annual growth rate of Shenzhen, Guangzhou and Dongguan is only about 5%, but at the beginning of the 21st century, the above-mentioned cities have the largest immigrant population base. Meanwhile, the central and western regions experienced a severe population loss process. The migration scale of cities in Chongqing, Henan, Jiangsu, Sichuan, Hunan, and Hubei exceeded 1 million, with an average annual growth rate of over -10%. Therefore, we find that during 2000-2010, the population migrated from the central and western regions to the eastern coastal areas, and most of the migration destinations were regional central cities. The northwest and northeast are regulated by national policies and have also become migration destinations for specific populations.

However, during 2010-2018, the spatial pattern of population migration has changed significantly. From Fig. 3-8 and Fig. 3-9, we find that: (1) Although the megacities in coastal areas are still the main destinations of population migration, such as Shanghai, Tianjin, and Guangzhou, their migration scale and average annual growth rate show an obvious downward trend, indicating that the immigration population in these areas is gradually slowing down. (2) The provincial capital cities and regional central cities in the central and western regions have become new migration destinations, such as Wuhan, Kunming, Zhengzhou, Chongqing and Changsha, which indicates that the direction of population migration has begun to transfer to the central and western regions. (3) The third-tier cities around the big cities in eastern China, such as Shandong, Fujian and Jiangsu, have experienced a serious process of population loss. 4) The scale of population migration in central and western hinterland cities such as Dazhou, Bazhong, Guangyuan, Anshan, Huanggang, and Qiqihar, Fushun, Anshan in Northeast China has steadily increased. (5) The continuous emigration of population in cities of Guangxi and Henan has become the most serious area of population outflow in China. In the long run, it will have an impact on the sustainable development of the region and form a “shrinking cities”.

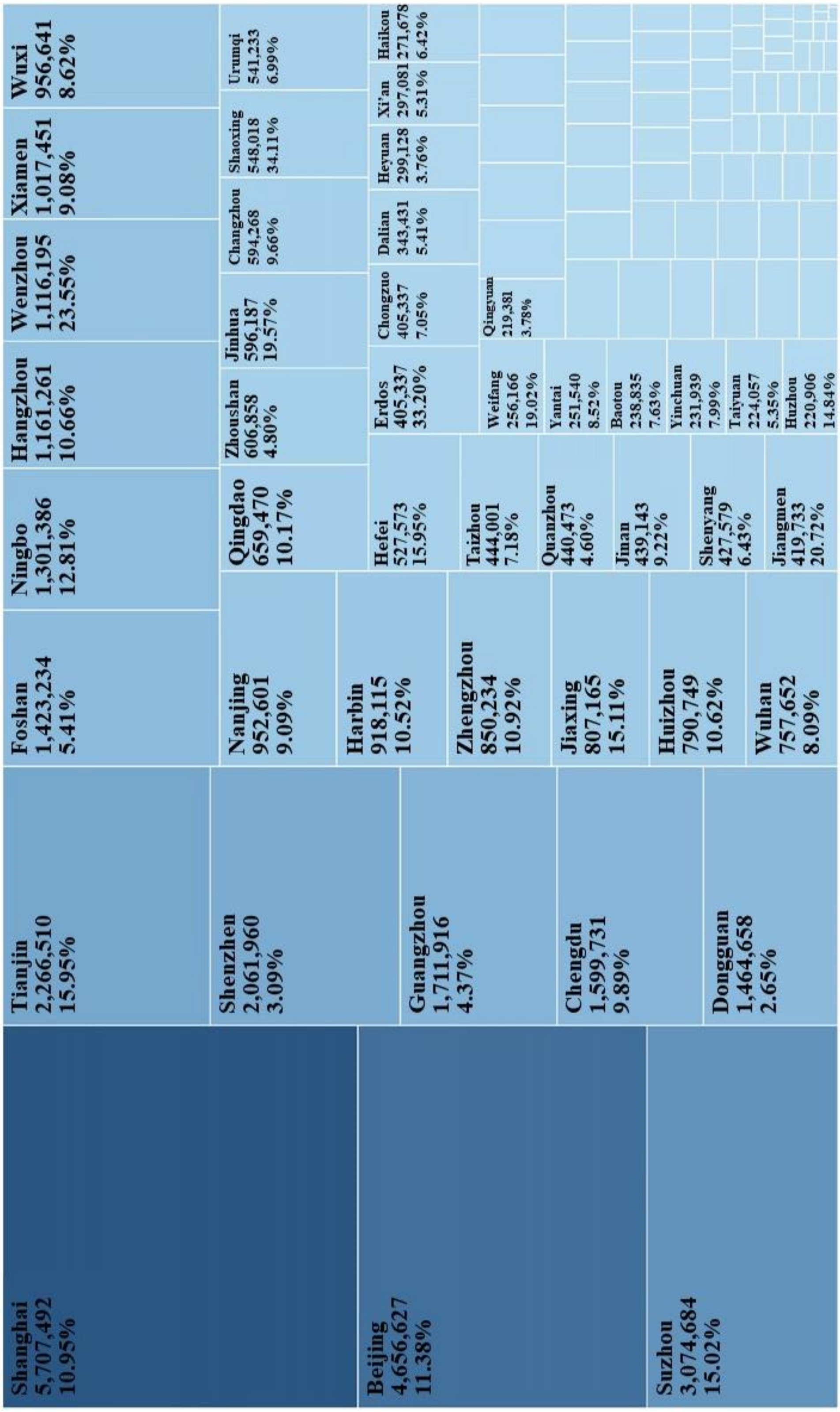


Fig. 3-6 Scale and average annual growth rate of population immigration in China's prefecture-level cities from 2000 to 2010.

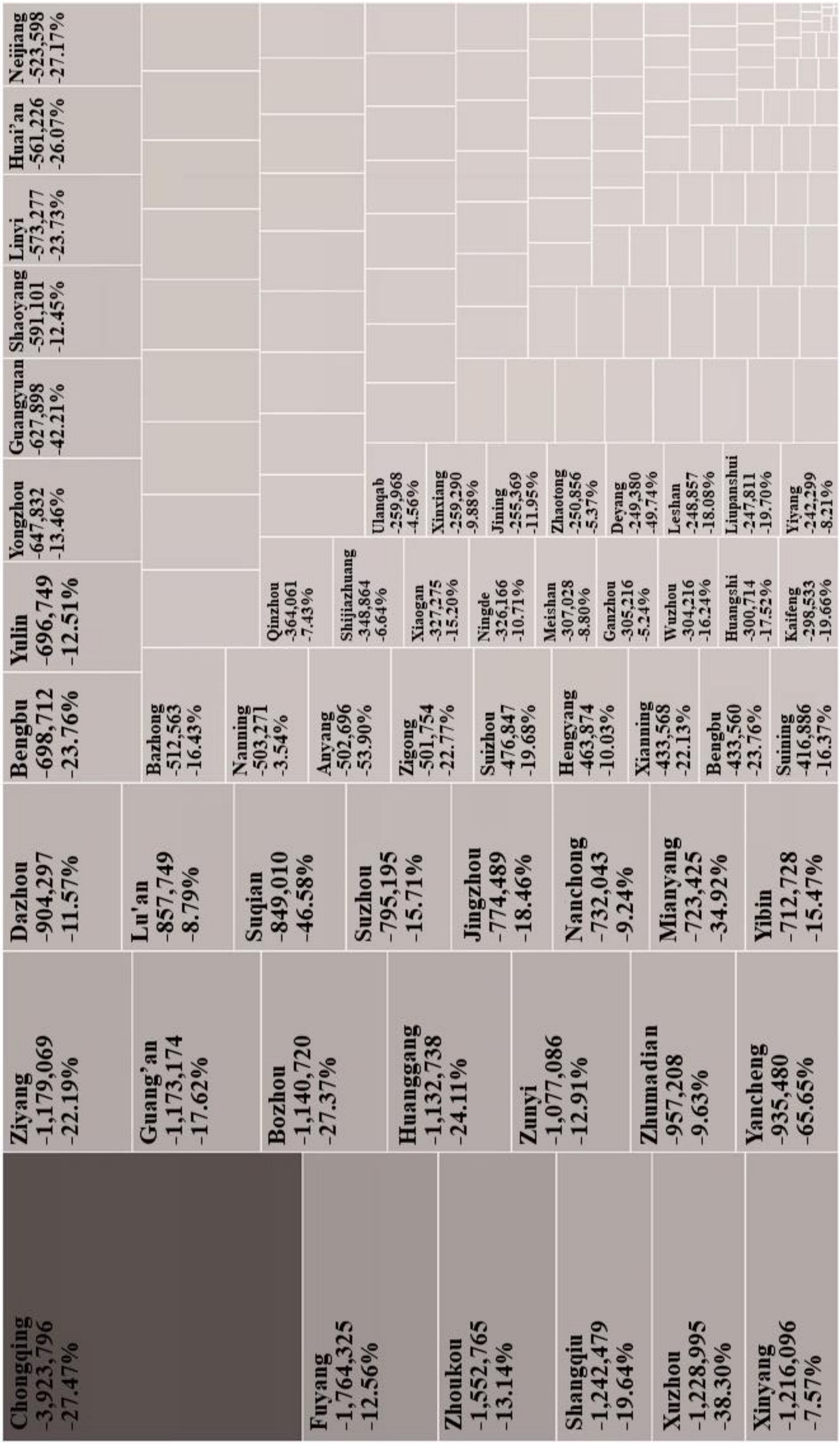


Fig. 3-7 Scale and average annual growth rate of population emigration in China's prefecture-level cities from 2000 to 2010.

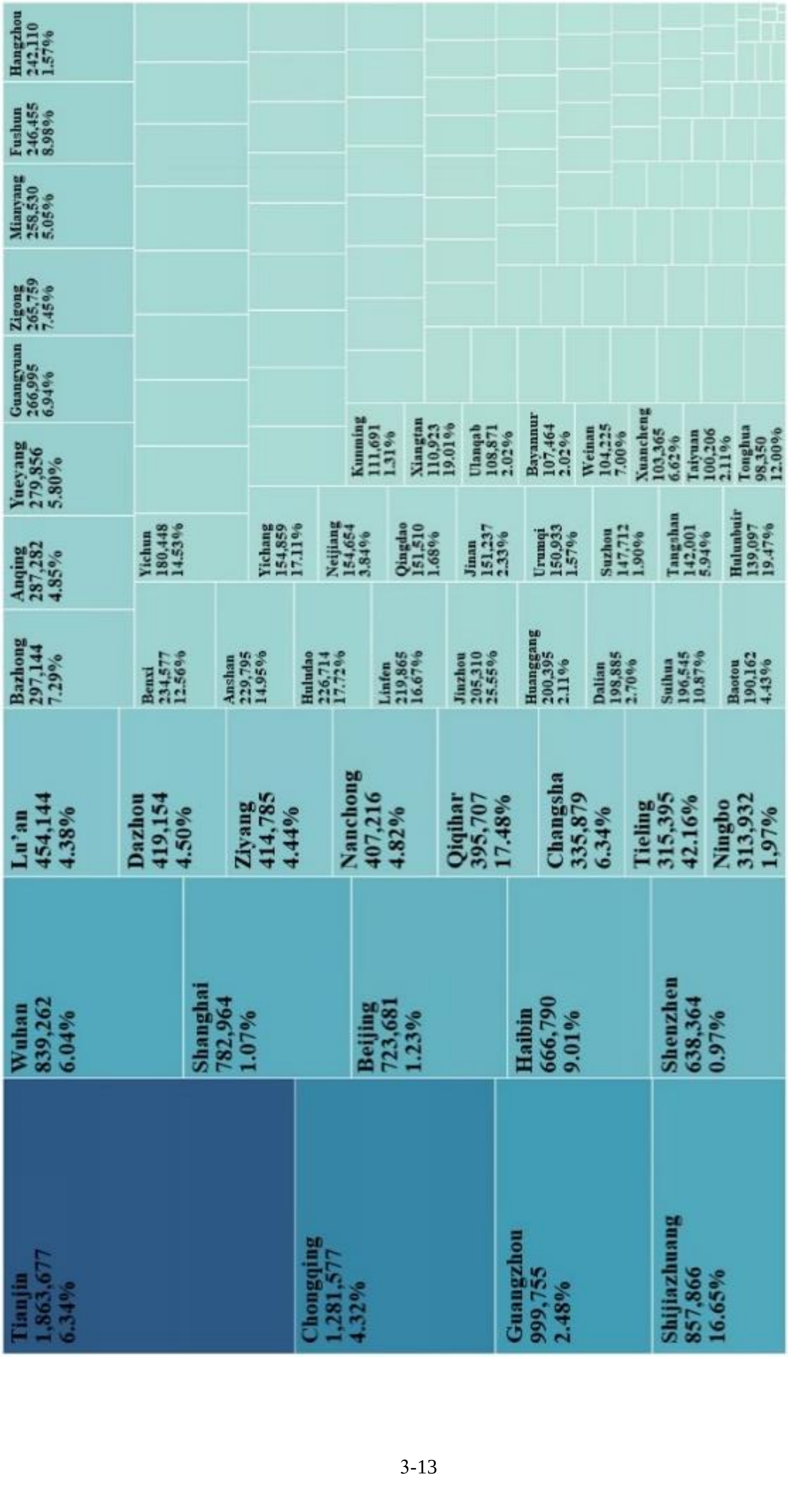


Fig. 3-8 Scale and average annual growth rate of population immigration in China's prefecture-level cities from 2010 to 2018.

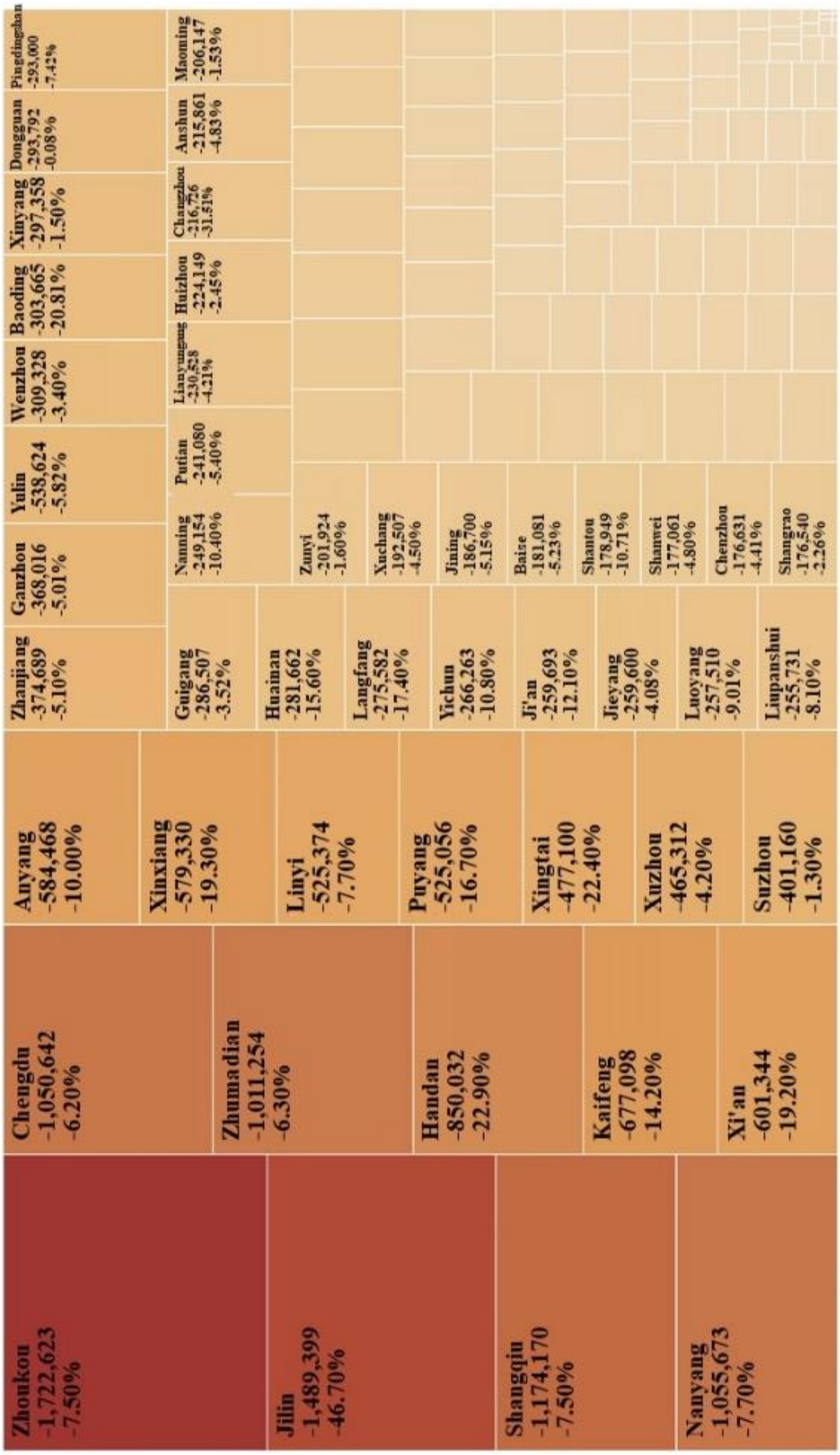


Fig. 3-9 Scale and average annual growth rate of population emigration in China's prefecture-level cities from 2010 to 2018.

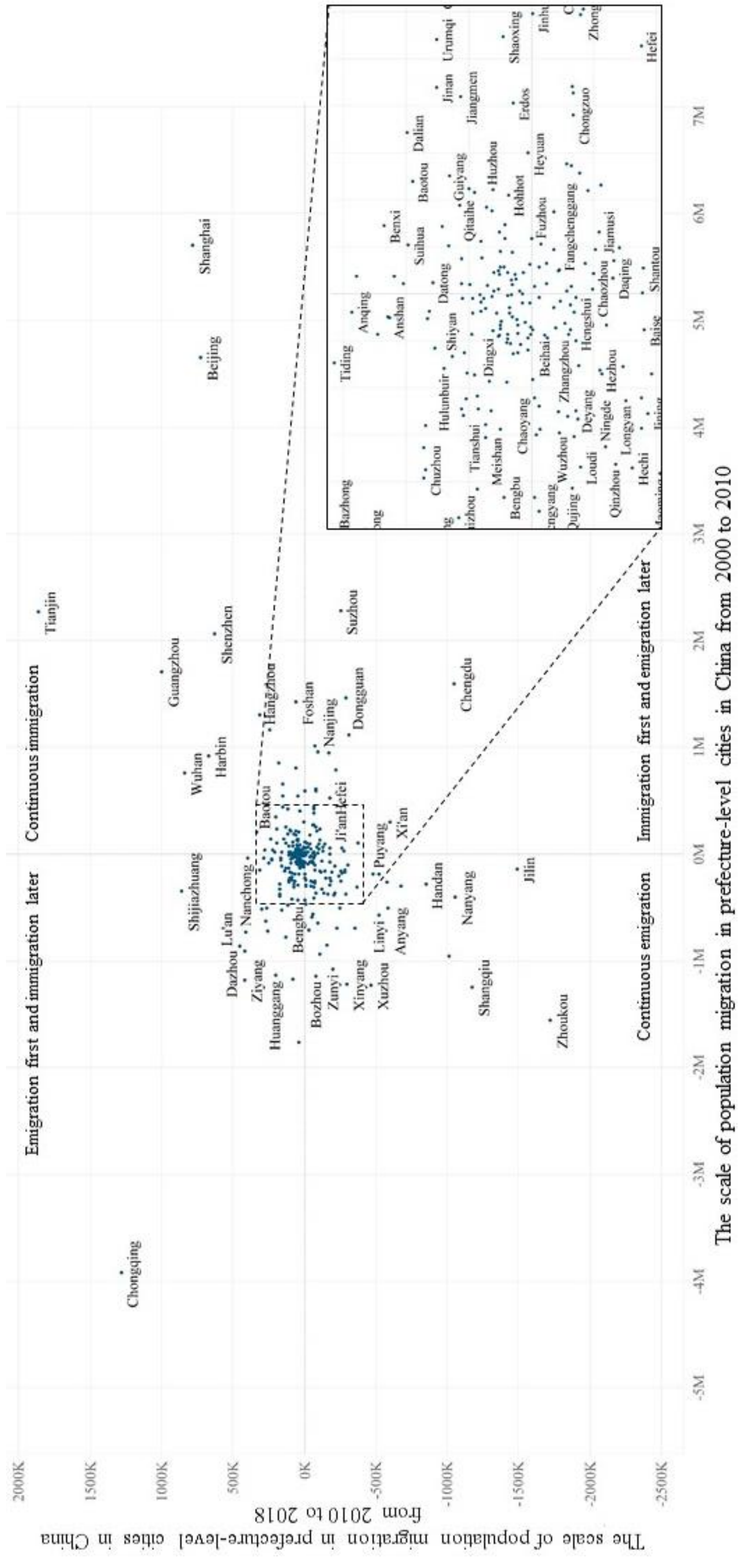


Fig. 3-10 Changes in population migration of prefecture-level cities in China from 2000 to 2010 and 2010 to 2018.

To further understand the changes in the scale of migrant population, we plot the bivariate relationship about the changes in the scale of population migration between 2000-2010 and 2010-2018 (see Fig. 3-10). According to the changes in the migration population in different cities, we divide into four types: continuous immigration (CI), continuous emigration (CE), immigration first emigration later (IE), emigration first and immigration later (EI). Table 3-3 lists four different types of cities. We can find that the megacities in the southeast coast and specific cities in the northwest and northeast regions belong to the CI type, the migrant population continues to increase. Cities in southwestern Shandong, central China, and southwestern China belong to the CE type, and the population continues to be lost, which is consistent with our previous analysis. With the transformation of China's economy and the intensification of the siphon effect of regional central cities, the migration population of the second and third-tier cities in the east and the surrounding cities has the process of increased first and then decreased sharply, such as Zhongshan, Zhanjiang, Zhaoqing and Shanwei in the Pearl River Delta region. Significant population changes, especially population loss, should be brought to the attention of city managers. The hinterland cities of the second and third-tiers in the central and western regions have experienced a process of first decreasing and then slowly increasing in the migration population, which also reflects the changes in the direction of population migration in the new era.

Table 3-3. Types of cities based on population migration.

Migration type	**Number**	**Cities**
Continuous immigration (CI)	65	Beijing, Shanghai, Tianjin, Shenzhen, Guangzhou, Foshan, Ningbo, Hangzhou, Wuhan, Qingdao, Harbin, Changsha, Kunming, Guiyang, Zhuhai and 50 other cities
Continuous emigration (CE)	100	Zhumadian, Zunyi, Yancheng, Laibin, Wuzhou, Tongling, Xinxiang, Xinyu, Anshun, Bozhou, Cangzhou, Ganzhou, Guigang, Hechi, Hengyang, Hengshui and 66 other cities
Emigration first and immigration later (EI)	89	Ankang, Anshan, Baicheng, Dezhou, Fuxin, Guangyuan, Hanzhong, Huaibei, Meishan, Neijiang, Sanmenxia, Shizuishan, Songyuan, Suining, Yan'an and 74 other cities
Immigration first and emigration later (IE)	46	Weifang, Binzhou, Changzhou, Chengdu, Dongguan, Huizhou, Jiamusi, Jinzhong, Langfang, Qingyuan, Shantou, Xi'an, Yangjiang, Yinchuan and 32 other cities

Through Moran's I statistics, we find that the patterns of population migration are spatially autocorrelated. Fig. 3-11 and 3-12 reports changes in the GMI from 2000-2010 and 2010-2018, which are significantly positive. From 2000 to 2018, the GMI shows a downward trend, indicating that the choice of migration destinations is diversified and no longer limited to a few core areas. In

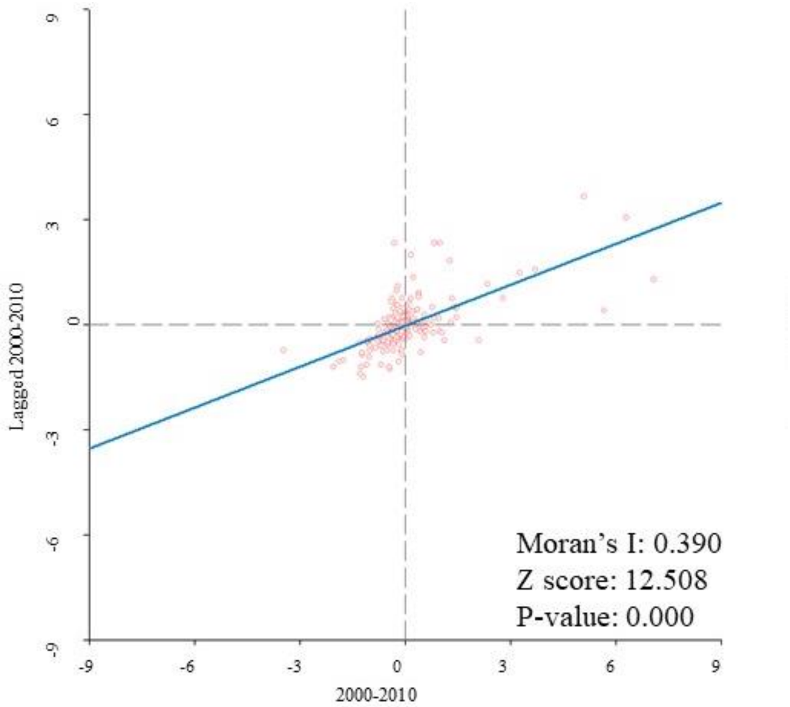


Fig. 3-11 Global Moran'I statistics from 2000 to 2010.

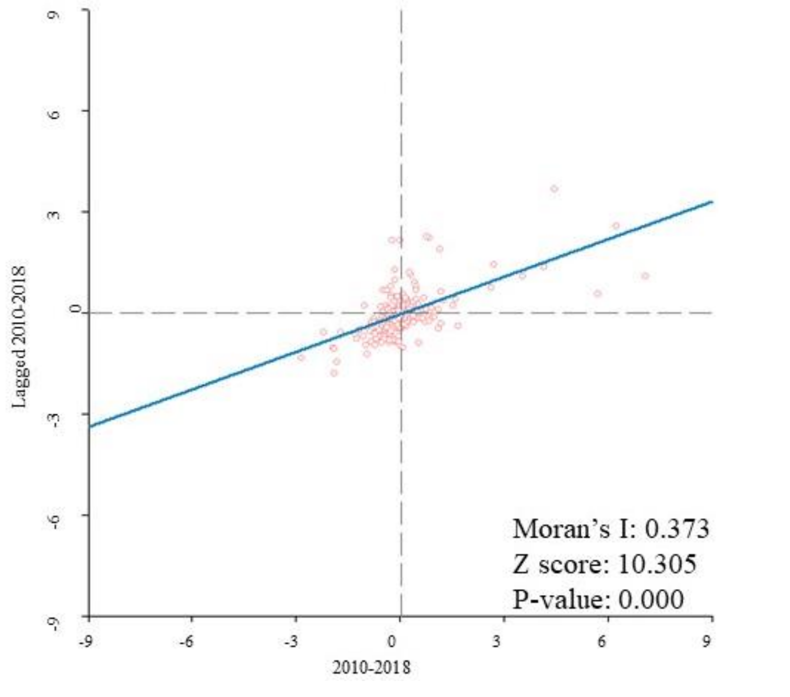


Fig. 3-12 Global Moran' I statistics from 2010-2018.

other words, in the context of China's new urbanization, the population is migrating towards more cities and regions. Fig. 3-13 and 3-14 reports the analysis results of LMI. From 2000 to 2010, the Yangtze River Delta (YRD), the Pearl River Delta (PRD) and the Beijing-Tianjin-Hebei (BTH) region were the most popular migration destinations, while the central area was the region with the most serious population loss. From 2010 to 2018, this spatial pattern has changed significantly. The choice of migration destination began to shift from coastal areas to inland areas, and Chongqing became the most popular destination. At the same time, Henan central China is still in a state of continuous population loss.

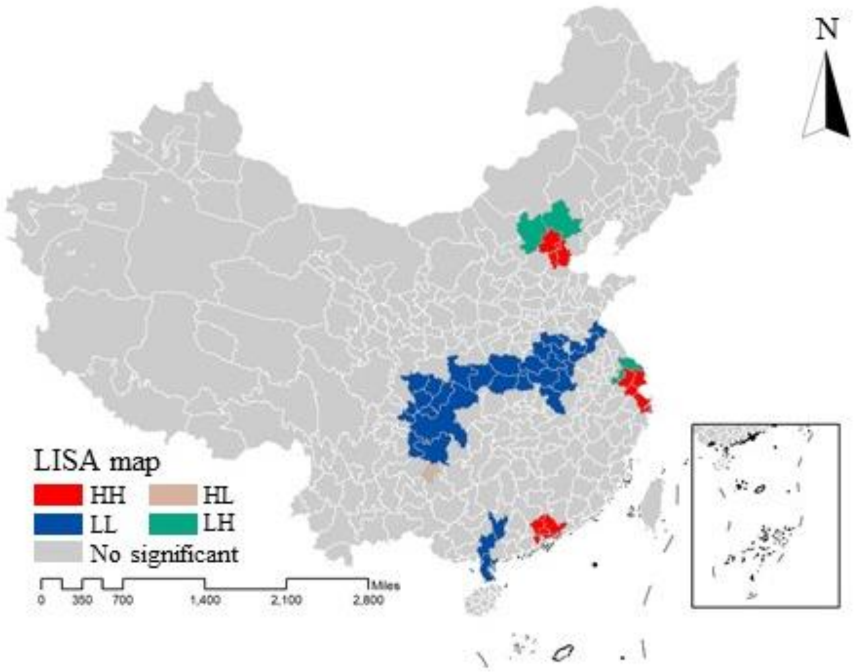


Fig. 3-13 Local Moran'I statistics from 2000 to 2010.

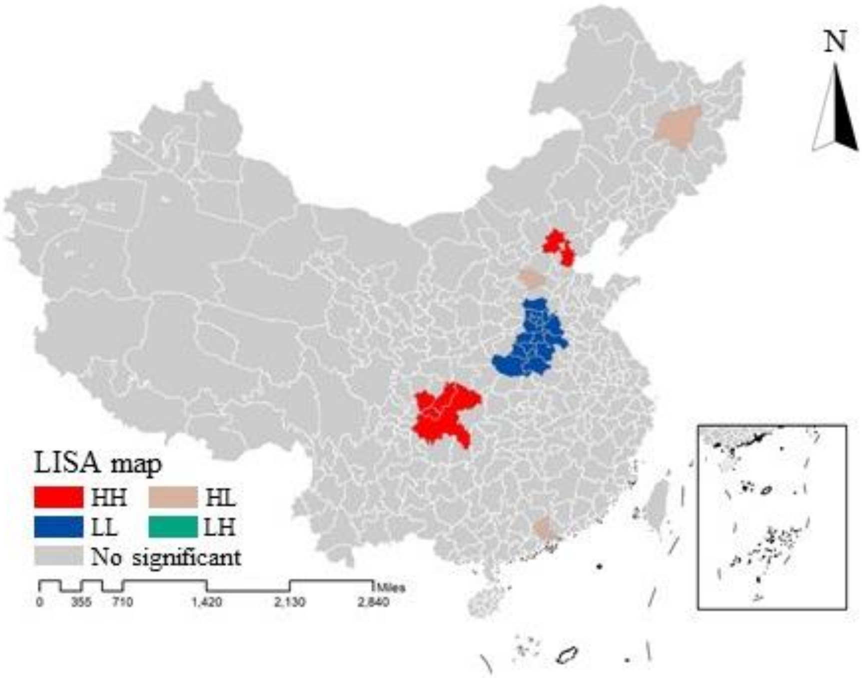


Fig. 3-14 Local Moran'I statistics from 2010 to 2018.

3.3.2 Changes in the driving force of population migration

Table 3-4 summarizes the basic parameters of the OLS, GWR and SGWR model outputs. We can see that the constructed SGWR model made a significant improvement over the normal regression model. Compared with the traditional regression model, the SGWR model had a smaller AICc and residual squares, larger adjusted R^2. In addition, the standard residual of the model was analyzed and the presented a random distribution pattern in space, indicating that the constructed SGWR model had better performance.

Table 3-4. Accuracy evaluation for the OLS, GWR and SGWR model.

Parameters	2000-2010			2010-2018		
	OLS	**GWR**	**SGWR**	**OLS**	**GWR**	**SGWR**
AICc	328.163	209.538	123.511	377.199	258.407	201.476
Adjusted R^2	0.727	0.871	0.915	0.795	0.856	0.923
Residual squares	45.727	20.125	8.713	54.412	37.108	15.969
Log-likelihood	-127.696	-85.249	62.412	-168.149	-66.218	4.704
Bandwidth	-	306	201	-	290	166

Table 3-5 and Table 3-6 statistics the output results of the global regression model in 2000-2010 and 2010-2018, respectively. From 2000 to 2010, seven variables including wage level, industrial upgrading, transport, education, life, health and housing price significantly affected population migration. The influence of other variables is not significant, so we will not discuss further. In 2010-2018, six variables including wage level, transport, health, air pollution, government financial pressure and migration control policies are significantly related to population migration, while other variables are not significant. Next, we will focus on the spatial and temporal changes in the driving forces of population migration from 2010 to 2018.

Table 3-5. Summary of the global regression model output from 2000 to 2010.

Variables	Estimate	SE	pseudo t	p-value
Intercept	0.000	0.032	0.000	-
AUHP	0.142	0.072	1.972	$p<0.05$
LIFE	0.079	0.040	1.984	$p<0.05$
GRLS	0.034	0.040	0.844	$p>0.05$
EDUC	0.150	0.042	3.595	$p<0.05$
HEAL	0.260	0.052	4.970	$p<0.05$
INFR	-0.096	0.052	-1.840	$p>0.05$
LOGE	0.123	0.065	1.909	$p>0.05$
GOFP	-0.104	0.055	-1.904	$p>0.05$

GOII	-0.011	0.059	-0.189	p>0.05
WAGE	0.183	0.066	2.770	p<0.05
GLOB	-0.067	0.062	-1.069	p>0.05
TRAN	0.121	0.047	2.584	p<0.05
URBA	-0.029	0.045	-0.637	p>0.05
INUP	0.145	0.053	2.721	p<0.05

Table 3-6. Summary of the global regression model output from 2010 to 2018.

Variables	Estimate	SE	pseudo t	p-value
Intercept	0.000	0.032	0.000	-
AUHP	-0.002	0.057	-0.030	p>0.05
MICP	0.111	0.047	2.369	p<0.05
LIFE	-0.018	0.035	-0.515	p>0.05
GRLS	0.001	0.056	0.020	p>0.05
ROAD	-0.029	0.057	-0.513	p>0.05
EDUC	-0.011	0.038	-0.280	p>0.05
HEAL	0.358	0.066	5.409	p<0.05
INFR	0.016	0.042	0.368	p>0.05
AIRP	-0.084	0.033	-2.524	p<0.05
LOGE	-0.013	0.035	-0.387	p>0.05
GOFP	-0.527	0.048	-10.969	p<0.05
GOII	-0.030	0.038	-0.770	p>0.05
WAGE	0.150	0.049	3.061	p<0.05
GLOB	-0.047	0.043	-1.084	p>0.05
TRAN	0.155	0.050	3.107	p<0.05

From 2000 to 2010, the development of social economy and the change of inhabitant environment have become the main driving force of population migration. The results of global regression show that there is a significant positive correlation between wage level and population migration, which indicates that high wage cities are attractive to population migration. In terms of living environment, such as life, education, health and transport, there is also a significant positive correlation on population migration, which indicates that the migrant population also pays attention to the development of urban public services. Although the level of public service in Chinese cities is improving, high-quality public resources are still concentrated in big cities. During this period, industrial upgrading also affected the population migration among regions. The estimated coefficient is positive, which shows that in cities where the economy mainly depends on service industry, the scale of population migration is increasing, because the large-scale service industry often has a large demand for labor, and the gap between jobs and income makes these cities attract a large number of populations. At the same time, we found that the estimated coefficient of housing prices is also significantly positive, indicating that rising housing prices can also bring a part of the migrant population. The possible explanation for this phenomenon is that the migration of people from rural to urban areas has led to an increase in housing demand and real estate investment in local cities.

In 2010-2018, the driving force of population migration has undergone major changes. First, as expected, differences in wage levels and economic development are still the main driving forces of population migration in China, but the estimated coefficient is smaller, indicating that the impact of wage levels on population migration is weakening. The narrowing of the economic gap between cities and the rising cost of living in big cities seem to make the migrant population no longer consider the wage level so important. Second, the estimated coefficient of inhabitant environment has changed most. The estimated coefficients of health and transport have been greatly improved, and air pollution has also significantly affected population migration. This means that migrants pay more attention to urban public services and living environment when choosing destinations. After 2010, as China has become an aging country and the age of the working population has increased, medical security has become the focus of consideration for migrants. The equalization and popularization of urban public resources can increase the sense of belonging of the migrant population, thereby attracting population inflows [19, 33, 34]. Finally, national policies and local financial pressures also have an impact on population migration during this period. The estimated coefficient of local fiscal pressure is significantly negative, indicating that the population is moving away from cities with high fiscal pressure. The estimated coefficient of the migration control policy is positive, showing a positive impact on population migration. We will focus on the analysis later.

Figs. 3-15 to 3-21 visualize the spatial changes and estimated coefficients of explanatory variables from 2010 to 2018. Fig. 3-11 to 3-14 shows that there are large spatial differences in the local R^2, indicating that as the urban spatial location changes, the explanatory variables have different explanatory powers to the dependent variables, which further reflects the spatial non-stationarity between the variables. These results are basically in line with our expectations, as explained in the following.

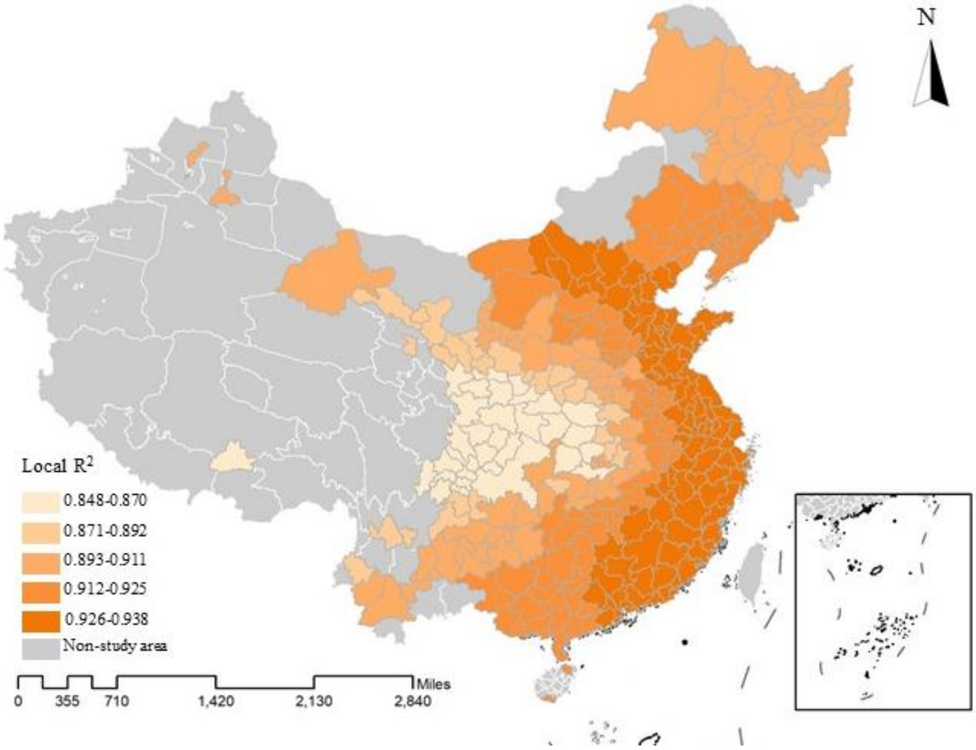


Fig. 3-15 Local R^2 based on SGWR model from 2010 to 2018.

In the global regression results, the coefficient of wage level is significantly positive, indicating that the scale of the population migration is positively correlated with the development of the local economy. As explained by neoclassical theorists, the income level of the intended destination is the main driving force of the migration process. With other costs unchanged and incomes increasing, more laborers will choose cities with higher salaries. This is consistent with the previous analysis results.

Changes in inhabitant environment also have a significant impact on population migration. Among them, health and transport have a significant positive correlation, and air pollution has a significant negative correlation, which indicates that the selection of migration destinations pays more attention to areas and cities with perfect social welfare and public services. The negative correlation of air pollution indicates that the population is fleeing from cities with poor living environment, which is inconsistent with the analysis and interpretation of previous study. We believe that under the conditions of regional economic improvement in the future, the preference of the migrant population will be more inclined to the urban public resources and living environment, and the rational allocation of urban resources has an important impact on attracting population inflows.

Government financial pressure has the strongest negative correlation, indicating that as local financial pressure increases, the scale of population migration will decrease. This is because local finance plays a vital role in economic and social development. The local fiscal deficit hinders investment in urban development and affects the normal life of urban residents. What we did not expect is that population migration control policies have a significant positive correlation, indicating that the population of cities with strong immigration control policies is still increasing. For this finding, a reasonable explanation is that strict population migration control policies have driven away non-skilled and low-educated immigrants, while having less impact on high-educated or skilled immigrants. The vacancy of jobs and the rise of wages make them choose the first-tier cities as the migration destination, which leads to the continuous immigration of population. In the long run, with the popularization and improvement of education, these policies may become less effective.

In the above paragraph, we explained the explanatory variables related to population migration. Considering that the model takes into account the non-stationarity of space, we will focus on explaining the variation of variables in space as follows.

From Fig. 3-16, we can see that wage level is positively correlated with population migration in all sample cities. The coefficient of all cities in southern China is higher than that of the north, which means that the wage of employees in the southern region is more closely related to population migration. The central and west regions such as Shaanxi, Sichuan, Qinghai, and Gansu have a weak positive correlation coefficient, which may be due to the fact that local wage growth still has not met the expectations of employees.

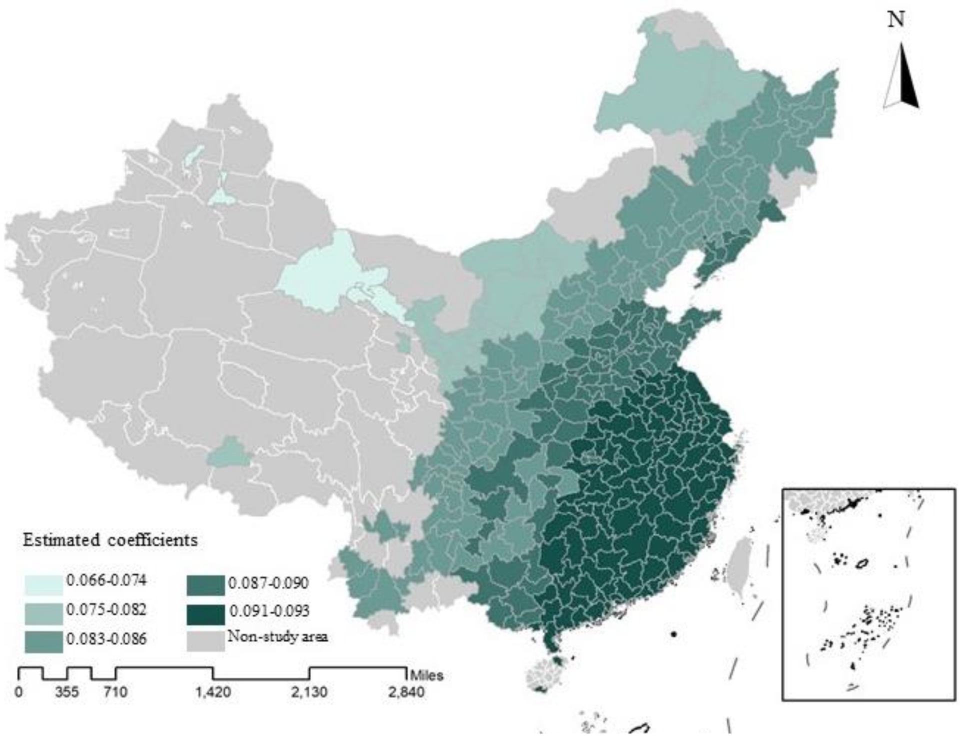


Fig. 3-16 Local R^2 of wage level.

In Fig. 3-17, we find the spatial variation of the estimated coefficients similar to that of Fig. 3-16. Compared with the north, cities in southern China are more able to attract population immigration in terms of improving transportation convenience. However, the spatial variation of this estimated coefficient is not obvious, and the spatial difference is small.

The estimated coefficient of health has the most significant spatial variation. In Fig. 3-18, we find that central China, especially provinces such as Henan, Hunan, Chongqing, and Anhui, have the strongest positive correlation coefficients. This shows that the above regions can reduce population

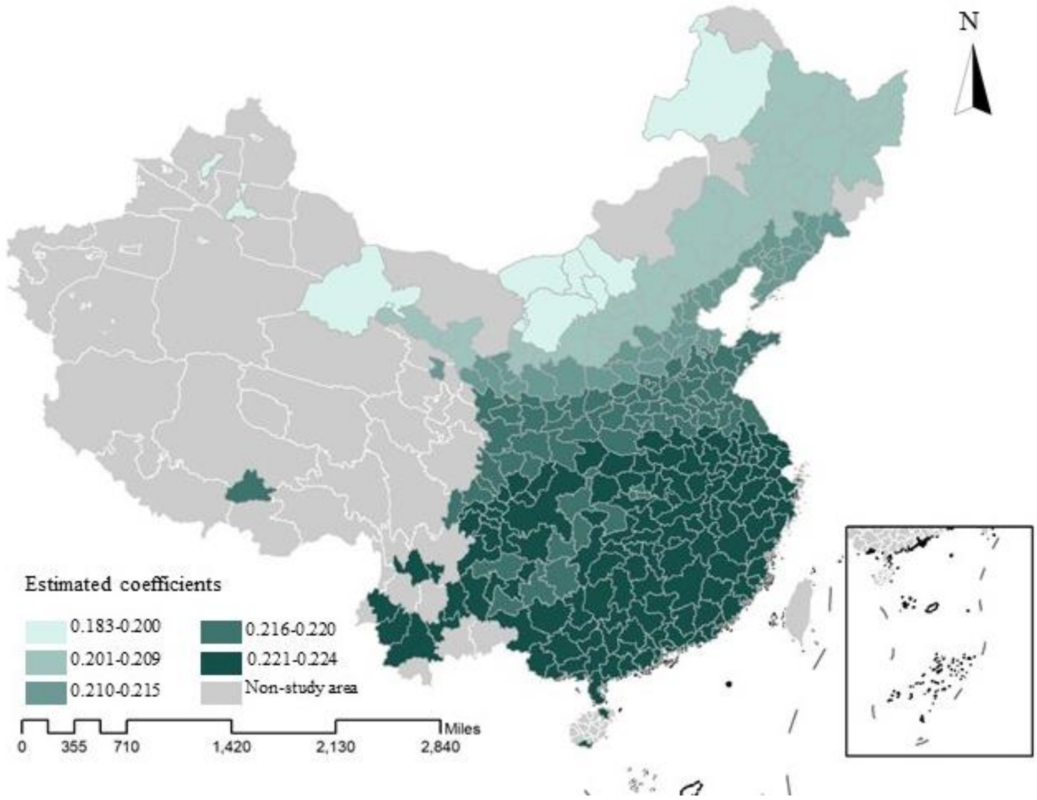


Fig. 3-17 Local R^2 of transport.

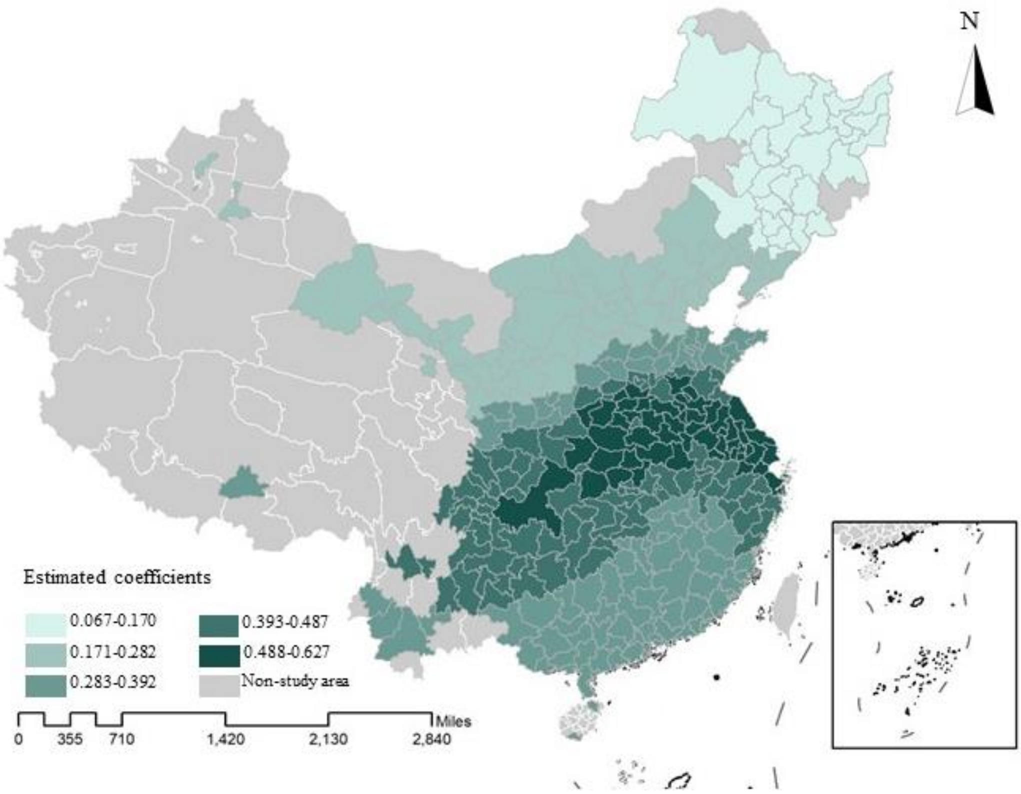


Fig. 3-18 Local R^2 of health.

emigration and even attract population immigration by improving medical conditions, especially the continuous population loss problem faced by some cities in Henan and Anhui. Based on the improvement of economic conditions in the future, paying attention to the construction of urban infrastructure, especially the popularization rate of medical resources, can effectively alleviate the problem of continuous population loss.

Air pollution has a significant negative impact on population migration in all sample cities (see Fig. 3-19). From a spatial point of view, there is an increasing trend from south to north, especially the Northeast region has the largest estimated coefficient. This is the result of the comprehensive effect of meteorological conditions, terrain factors and industrial development. Although this estimated coefficient is not the largest, but in the future when the gap in economic conditions narrows, the inhabitant environment will become an important choice for migration destinations, and air pollution will also be considered by migrant population.

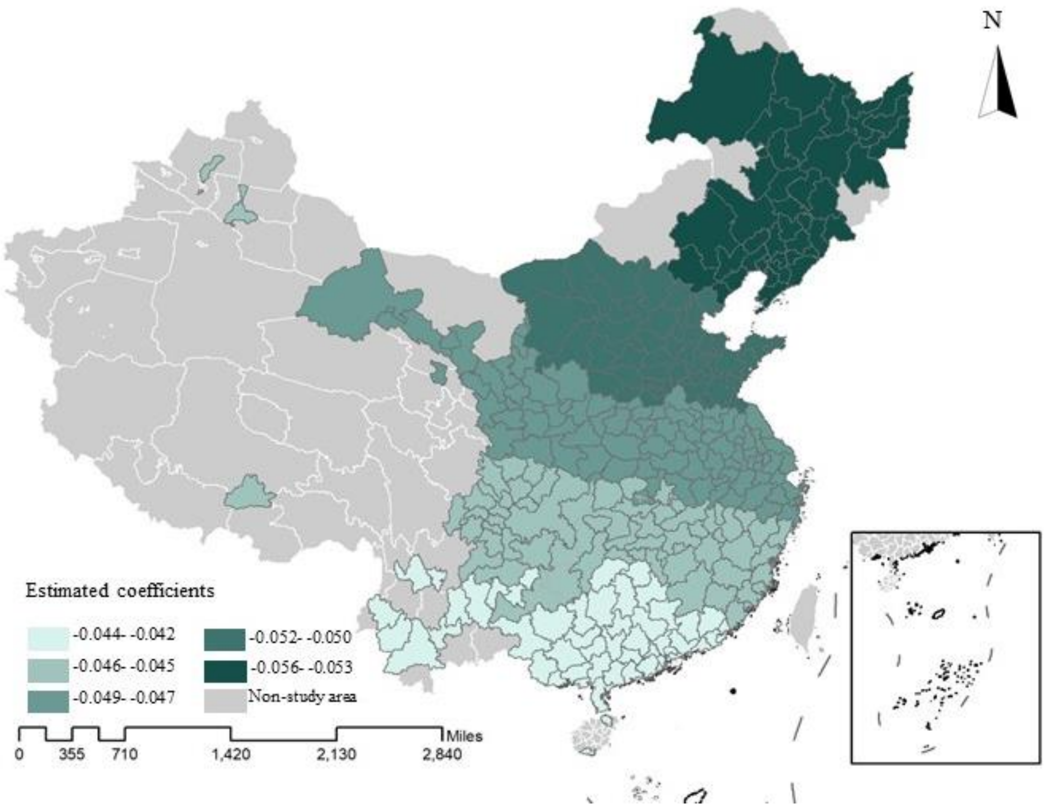


Fig. 3-19 Local R^2 of air pollution.

Government financial pressure also has a significant negative impact on population migration in all sample cities (see Fig. 3-20). The central and western regions such as Yunnan, Chongqing, Shaanxi, Gansu and Western Henan have the largest estimated coefficients, indicating that the impact of fiscal pressure on population migration in these regions is more obvious than in other regions. The estimation coefficient of southeast coastal cities is the smallest. This is closely related to the development of local economy. The financial pressure of coastal cities such as Shanghai, Guangzhou and Shenzhen are also large, but the strong local economy ensures the financial revenue and expenditure operation. Although it still has a negative impact on population migration, the impact is almost negligible. On the contrary, the hinterland cities in the central and western regions have a single economic structure and a large population. For example, the permanent resident

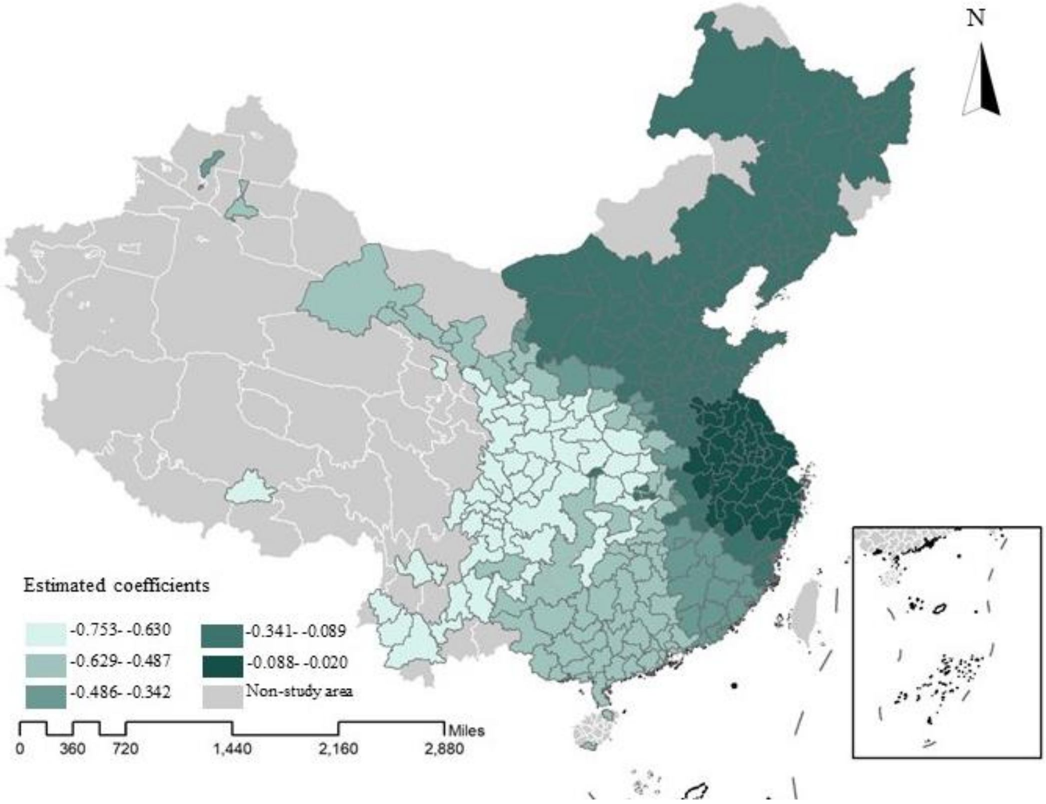


Fig. 3-20 Local R^2 of government financial pressure.

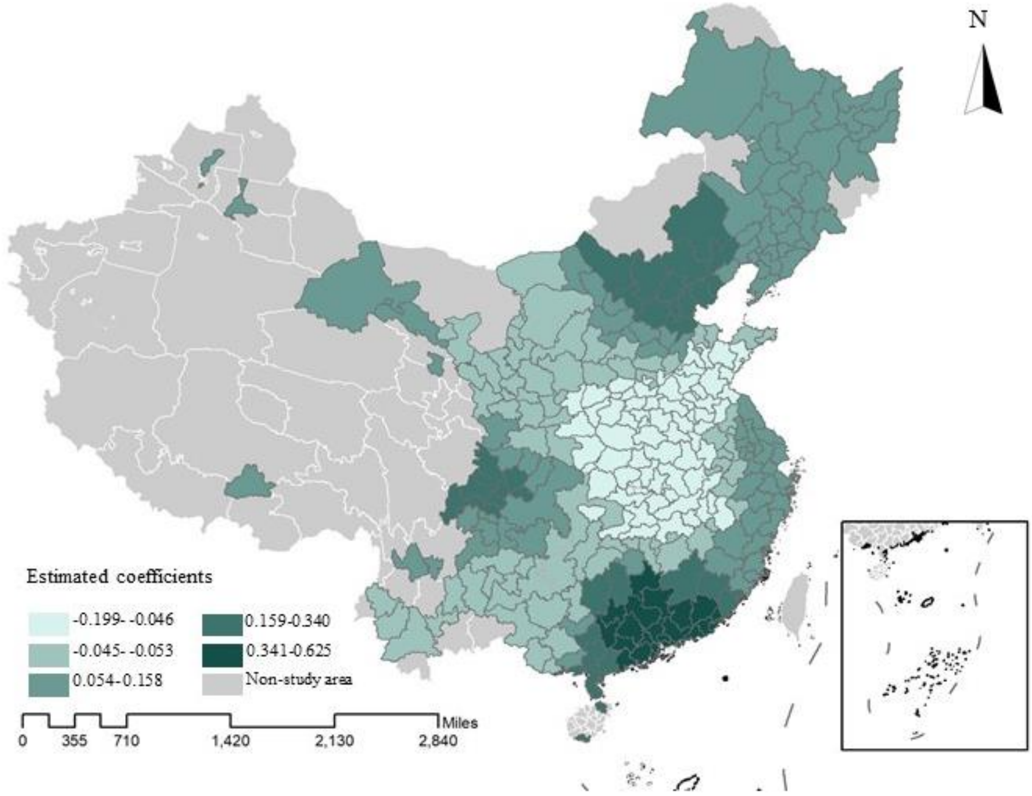


Fig. 3-21 Local R^2 of migration control policy.

population of Chongqing has exceeded 30 million. The large number of public resources invested makes the local finance increasingly tense. In the long run, it will fall into a vicious circle, which is unfavorable to economic development and population migration.

Fig. 3-21 shows that there are both positive and negative correlations between migration control policies and population migration. It has the largest positive correlation coefficient in the Pearl River Delta (PRD) and the Beijing-Tianjin-Hebei (BTH) region, which means that with the implementation of policies, the above-mentioned regions can still attract population inflows. The possible reason is that as high-end industries replace labor-intensive industries, cities such as Beijing, Guangzhou, and Shenzhen are driving away unskilled or low-educated populations and attracting highly educated and skilled talents. The increase in job opportunities and the guarantee of income make the population control policy have little impact on them. In the central region, the estimated coefficient is significantly negative, indicating that the implementation of policies has a restraining effect on population migration. For example, Jinan and Qingdao, although there are still population inflows, the scale and growth rate show an obvious downward trend. Therefore, population control policy will continue to affect population migration in the future, but the gap between cities will become more obvious.

3.4 Summary

Based on the data from the fifth and sixth census and the statistical yearbook, this chapter explores the spatial and temporal changes of China's population migration pattern and its driving forces, analyzes the impact of socioeconomic, Inhabitant environment and public policies on population migration across the country. This study uses exploratory spatial data analysis methods to clearly depict the spatial and temporal patterns of population migration, constructs a semiparametric geographically weighted regression model to observe the spatial non-stationarity and spatial heterogeneity of driving factors, with a particular focus on the population migration after 2010, thus contributing to the existing literature.

This research fins that during 2000-2010, the destinations of population migration were concentrated in the Pearl River Delta, the Yangtze River Delta, and the Beijing-Tianjin region, while the inland cities in the central and western regions were lost due to large numbers of population outflows. Our findings are consistent with the literature. The imbalance of regional economic development is the main force of population migration, but at the same time, the popularization of urban public services also attracts laborers from poor areas.

However, after 2010, the choice of destination of migrant population has changed significantly. The scale of population migration in the eastern coastal mega cities is gradually slowing down, while the phenomenon of population emigration even appears in the surrounding second and third-tier cities. Meanwhile, the central and western inland capital cities and regional central cities have large-scale population immigration, the population immigration in the second and third-tier hinterland cities is also obvious. This shows that the spatial structure of China's population migration has been reversed since 2010. After 2010, although economic level is still the main driving force of population migration, its impact is becoming smaller. On the contrary, the migrant population seems to pay more attention to the equalization and popularization of public services in the migration destination. The satisfaction of living facilities seems to be more attractive than economic development, especially in Central China, such as Chongqing, Henan, Anhui and Hubei.

We also have some unexpected findings. Strict population control policies do not seem to have a negative impact on population migration in megacities such as Beijing, Tianjin, Shanghai, Guangzhou and Shenzhen. In the local regression analysis, the estimated coefficients of the above cities are significantly positive, indicating that the implementation of the policy has promoted the migration of urban population. The possible reason is that the policy has driven away low-educated and low-skilled labor and attracted highly-educated and highly skilled talents. In the long run, the development gap between cities will become larger, and these policies may become less effective in the future.

A better understanding of the spatial and temporal changes of population migration patterns and its driving forces can provide valuable reference for the government and planners to formulate scientific urban development policies. Population is a barometer of the rise and fall of a city. The loss of population not only makes the city lack of vitality, it is more likely to cause the city to shrink. In the context of the continued slowdown in population growth in the future, how to reduce the loss of local labor while absorbing migrants to settle down in the city is a challenge for urban managers and urban policymakers. With the relaxation of the policies for settlement in first-tier cities, how to improve the competitiveness of cities to attract population will become crucial. The improvement of urban infrastructure and the public services will enable more migrants to gain a sense of urban belonging, thereby attracting population to settle down, especially the middle-aged migrants and family migrants.

Although our research is of great significance, it also has some limitations. First, the National Development and Reform Commission (NDRC) of China issued the “Key Tasks for New Urbanization Construction in 2019” in 2019. In this report, the government emphasized to continue to increase the reform of the household registration system and promote removal of settlement restrictions in type II cities with a permanent urban population of 1-3 million. However, this part of the content was not within the time period of our research. With the seventh national census, we will make more detailed research in the future. Second, the impact mechanism of population migration is complex and comprehensive. This paper only discusses from three aspects of public policy, socioeconomic and inhabitant environment, the results may have certain deviation and one-sidedness. For example, although the population control policy has a positive impact on the population migration in Chengdu, the problem of continuous population loss has emerged since 2010, and it has become the only inland city with population emigration. In the next stage, we will combine micro-level population questionnaire survey data to analyze the impact mechanisms of population migration in different cities, such as family factors and individual factors.

Reference

1. Sun, M.; Fan, C. C., China's permanent and temporary migrants: Differentials and changes, 1990–2000. *The Professional Geographer* **2011,** 63, (1), 92-112.
2. Wang, Y.; Dong, L.; Liu, Y.; Huang, Z.; Liu, Y., Migration patterns in China extracted from mobile positioning data. *Habitat International* **2019,** 86, 71-80.
3. Chan, K. W., Migration and development in China: Trends, geography and current issues. *Migration and Development* **2012,** 1, (2), 187-205.
4. Liu, T.; Qi, Y.; Cao, G.; Liu, H., Spatial patterns, driving forces, and urbanization effects of China's internal migration: County-level analysis based on the 2000 and 2010 censuses. *Journal of Geographical Sciences* **2015,** 25, (2), 236-256.
5. Wang, X., Sustainability of China's economic growth. **1999**.
6. Wang, Z., Floating Population and Economic Development: A Study Based on Prefectural Data. *Modern Urban Research* **2013**, (3), 4.
7. Freeman, R. B., A labor market with Chinese characteristics. **2015**.
8. Cai, F.; Guo, Z.; Wang, M., New urbanisation as a driver of China's growth. *China's new sources of economic growth* **2016,** 1, 43-64.
9. Li, B.; Chen, C.; Hu, B., Governing urbanization and the new urbanization plan in China. *Environment and Urbanization* **2016,** 28, (2), 515-534.
10. NHFPC, The 2018 Report on China's Migrant Population Development. *Beijing: The State Council* **2018**.
11. Liu, Y.; Shen, J., Modelling skilled and less‐skilled interregional migrations in China, 2000–2005. *Population, Space and Place* **2017,** 23, (4), e2027.
12. Wu, J.; Wei, Y. D.; Li, Q.; Yuan, F., Economic transition and changing location of manufacturing industry in China: A study of the Yangtze River Delta. *Sustainability* **2018,** 10, (8), 2624.
13. Cao, J. C.; Zhang, Q., Impacts and Countermeasures of the Reform of Household Registration System on Family Planning. *Market & Demographic Analysis* **2002**.
14. Xie, B., Residence Points System:Another "Patch" of Household Registration Reform? Characteristics,Problems and Countermeasures of Shanghai Residence Points System. *Population Research* **2014**.
15. Liu, J., The effect, problems and countermeasures of the new round reform of household registration system (in Chinese). *Population and Society* **2018,** 34, (4), 89-98.
16. Chen, J.; Chen, S.; Landry, P. F., Migration, environmental hazards, and health outcomes in China. *Social Science & Medicine* **2013,** 80, 85-95.
17. Cheng, Z.; Nielsen, I.; Smyth, R., Access to social insurance in urban China: A comparative study of rural–urban and urban–urban migrants in Beijing. *Habitat International* **2014,** 41, 243-252.
18. Li, S.-M.; Chan, K. W.; He, S., Migration, mobility, and community change in Chinese cities: introducing the special issue. In Taylor & Francis: 2014.
19. Lin, L.; Zhu, Y.; Ke, W.; Wang, j., The impact of migrants' access to urban public services on their urban settlement intentions: A study from the perspective of different-sized cities. *Acta Geographica Sinica* **2019,** 74, (4), 737-752.
20. Zhou, Y., Macro effects on the household formation of China's young adults–demographics, institutional factors, and regional differences. *International Journal of Housing Policy* **2017,** 17, (4), 512-540.

21. Zhu, Y.; Chen, W., The settlement intention of China's floating population in the cities: Recent changes and multifaceted individual - level determinants. *Population, Space and Place* **2010,** 16, (4), 253-267.

22. Fan, X.; Liu, H.; Zhang, Z.; Zhang, J., The spatio-temporal characteristics and modeling research of inter-provincial migration in China. *Sustainability* **2018,** 10, (3), 618.

23. He, C.; Chen, T.; Mao, X.; Zhou, Y., Economic transition, urbanization and population redistribution in China. *Habitat International* **2016**, (51), 39-47.

24. Johnson, D. G., Provincial Migration in China in the 1990s. *China Economic Review* **2003,** 14, (1), 22-31.

25. Shen, J., Increasing internal migration in China from 1985 to 2005: Institutional versus economic drivers. *Habitat International* **2013,** 39, 1-7.

26. Yang, Z.; Gao, W.; Zhao, X.; Hao, C.; Xie, X., Spatiotemporal Patterns of Population Mobility and its Determinants in Chinese Cities Based on Travel Big Data. *Sustainability* **2020,** 12, (10), 4012.

27. Liu, T.; Qi, Y.; Cao, G., China's floating population in the 21st century: Uneven landscape, influencing factors, and effects on urbanization. *Acta Geographica Sinica* **2015,** 70, (4), 567-581.

28. Brunsdon, C.; Fotheringham, A. S.; Charlton, M. E., Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis* **1996,** 28, (4), 281-298.

29. Feuillet, T.; Salze, P.; Charreire, H.; Menai, M.; Enaux, C.; Perchoux, C.; Hess, F.; Kesse-Guyot, E.; Hercberg, S.; Simon, C., Built environment in local relation with walking: Why here and not there? *Journal of Transport & Health* **2016,** 3, (4), 500-512.

30. Lee, K. H.; Heo, J.; Jayaraman, R.; Dawson, S., Proximity to parks and natural areas as an environmental determinant to spatial disparities in obesity prevalence. *Applied geography* **2019,** 112, 102074.

31. Jin, C.; Xu, J.; Huang, Z., Spatiotemporal analysis of regional tourism development: A semiparametric Geographically Weighted Regression model approach. *Habitat International* **2019,** 87, 1-10.

32. Fotheringham, A. S.; Brunsdon, C.; Charlton, M., *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons: 2003.

33. Ye, P., Residential Preferences of Migrant Workers: An Analysis of the Empirical Survey Data from Seven Provinces/Districts [J]. *Chinese Journal of Sociology* **2011,** 2.

34. Lin, L.; Zhu, Y., Spatial variation and its determinants of migrants' hukou transfer intention of China's prefecture-and provincial-level cities: Evidence from the 2012 national migrant population dynamic monitoring survey. *Acta Geographica Sinica* **2016,** 71, (10), 1696-1709.

Chapter 4. Spatiotemporal patterns of population mobility and its determinants in urban China during Spring Festival of 2019

4.1 Introduction

Since the reform and opening up in 1978, the rapid economic development and the process of social modernization in China have made population mobility between cities more common. Population mobility is considered as the re-allocation of production factors in space; the mobility of population in a specific space promotes the reaggregation and diffusion of social and economic factors, thus reshaping patterns of population distribution [1]. From 2000 to 2010, China's floating population increased by 109% [2]. In 2016, China's floating population reached 245 million, accounting for 18% and 76% of the total population in China and the United States, respectively (according to Houkou system in China, the floating population refers to those who live away from their original Hukou system place and live for work). As China is one of the countries with the most extensive population mobility in the world, cross-regional and large-scale population mobility will become an important force to promote urbanization.

Many scholars have carried out in-depth studies on the origin, development, and internal mechanism of population mobility and migration, including the "immigration law", the "push and pull theory", the "macro-neoclassical theory", the "micro-neoclassical theory" and the "equilibrium law theory". Ravenstein's "immigration law" first proposed clear regulations related to population migration [3]. Bogue tried to explain the internal mechanism of population mobility by using the push-pull theory [4]. Macro-neoclassical theory holds that the income level of a future destination is the main force driving population mobility and migration [5, 6]. Micro-neoclassical theory abstracts the mobility and migration of the population as a form of capital investment to maximize individual interests [7]. Equilibrium theory argues that destination amenities (environment, services, and living) play a key role in population mobility and migration, rather than the expected difference in benefits between regions [8, 9].

Based on the above theories, China's population mobility and migration have been studied. Among them, Liu, Shen and Fan analyzed the spatial characteristics and patterns of migration [10, 11], Fang and Dewen explored the determinants of population migration [12], Gu, Wang and Yu studied the formation mechanism of the spatial pattern of the floating population [13-15], and Fan discussed the impact of population mobility and migration on regional development [10]. The above research methods are usually based on gravity models or push-pull theories. Meanwhile, differences in regional economic development and market forces are always considered to be key factors causing population mobility [10, 16]. Since the economic reform, China's coastal regions have rapidly improved the regional economy relying on location advantages and institutional advantages [17, 18]. This region has attracted a large number of populations under the influence of a significant increase in national productivity and the reform of the household registration system [19, 20]. With the implementation of investment environment and central policy, the central and western regions of China have accelerated the pace of urban development and played an increasingly important role in population mobility. Therefore, the analysis of the pattern of population mobility and the discussion of the role of socio-economic factors among different cities can help to understand demographic change and promote regional sustainable development.

However, previous studies on China's floating population have some limitations. On the one hand, all the studies are based on the census data; due to problems such as low data accuracy and long update time, they cannot describe the spatiotemporal patterns of population mobility in a timely, accurate and dynamic manner [21-24]. On the other hand, the existing research mainly concentrated on the Yangtze River Delta, the Pearl River Delta, the Beijing–Tianjin–Hebei urban agglomeration and megalopolis, whereas small and medium-sized cities are not often mentioned [25-28]. However, with the trend of inland development of China's economy, the central and western regions are increasingly becoming an important force to promote the development of national urbanization; therefore, the research on the above-mentioned cities are equally important.

In recent years, with the rapid development of 3S technology, comprehensive and continuous observation of human spatiotemporal behavior data including geographic location, social attributes, movement trajectories, migration processes, and interaction patterns has become possible. Urban research based on Sina Weibo's check-in data [29], public comment data [30], residents' geographic behavior data [31], bus credit card data [32], and traffic travel big data has become a reality. These methods are real-time, objective, easy to analyze, and predictable, thus making up for the shortcomings of traditional survey methods (such as questionnaire, sampling, and census data). Therefore, they can provide sufficient and more accurate real-time data.

Meanwhile, with the in-depth research in the field of sociology, some quantitative analysis methods such as the social network analysis method have been widely used. Social network analysis (SNA) is the process of investigating social structures through networks and graph theory [33]. It characterizes networked structures in terms of nodes and the ties, edges, or links that connect them. It is a complex network based on nodes and connections to measure and map various aspects or relationships among people, organizations and groups. Now, this method has emerged as a key technique in the modern social sciences—it is widely used in anthropology, biology, demography, communication studies, economics, geography, history, and information science—and has achieved remarkable results [34]. Therefore, it can provide sufficient technical support for a quantitative analysis of population mobility.

The Spring Festival is the biggest and most important festival in China. During the Spring Festival, there will be a large number of people traveling between their work place and hometown; this special phenomenon is called "ChunYun". According to the statistics from the Ministry of Transport of the People's Republic of China, the national transportation system had a total of 1.395 billion passengers during the 2019 "ChunYun", which provides us with a good opportunity to study the spatial and temporal distribution of population mobility [35, 36]. Compared with the tourism flow during the National day golden week, the population flow during the Spring Festival is characterized more by return flow and student flow. The national transportation system encounters enormous challenges during the Spring Festival travel rush and the various social problems that occur not only cause public concern, but also increase government spending. Therefore, it is of great practical significance to explore the factors influencing population mobility patterns related to it (unless otherwise noted, the Spring Festival referred to in this chapter is that of 2019).

Based on the above analysis, this chapter makes three main contributions:

1. Based on the dataset with high spatial and temporal resolution, a more accurate population mobility pattern was found during the Spring Festival, avoiding the spatial mismatch caused by the impact of the census dataset and the Hukou system. (The Hukou system is a special basic administrative system in China. It is a household management policy implemented by the People‘s Republic of China for its citizens. The Hukou system shows the legitimacy of the natural person‘s life in the local area. There are two main aspects of China’s Hukou system. First, agricultural and nonagricultural Hukou. Due to extensive disputes, the state has gradually cancelled the division of agricultural and nonagricultural household registration. The second is the location of the Hukou. When a person is born in a certain place, he or she will be automatically registered to the local Hukou. In this era of high population mobility, this system leads to a serious problem of “the separation of people and Hukou”, which leads to a series of political and social problems; this kind of population is called the floating population.)

2. Based in the changes of net population mobility in different cities before and after the Spring Festival, this paper described the population activities in prefecture-level cities in China and analyzed the characteristics of the population mobility network through the social network analysis method, including the spatial structure, with key cities as nodes and typical “small world” features.

3. Furthermore, the semiparametric geographically weighted regression (SGWR) model was applied to explore the determinants of population mobility from the perspective of urban development. The results showed that average wage, urbanization rate, foreign capital, value-added of primary industry, and value-added of secondary and tertiary industry are closely related to population mobility. When the spatial heterogeneity and nonstationary was considered, the socioeconomic factors that affect population mobility were different in different regions and cities. The spatial disparity of these social economic factors was further discussed and development strategies among cities were analyzed.

The research data and area are discussed in 4.2. 4.3 outlines the methods. 4.4 presents the results, including the spatial and temporal patterns of population mobility during the Spring Festival, and socioeconomic factors associated with population mobility. The summary is given in 4.5.

4.2 Study areas and data

Location-based services (LBS) obtain the geographical location of a mobile user through the wireless communication network or the external positioning method of network operators. When users allow various mobile applications to call LBS services, their movement trajectories will be accurately recorded in real-time through the positioning information. The movement of a single user in geographical space seems to be random, but may take on a specific pattern when a large population group is accessed. According to the statistical report on the development of the Internet in China, by the end of 2018, the number of instant messaging users had reached 792 million, and the number of mobile Internet users had exceeded 817 million, accounting for 98.6% of Internet users using mobile phones [37]. In this context, every smartphone user can be seen as a mobile sensor, reflecting social characteristics and allowing for the collection of a massive amount of individual movement information, in real time and in an efficient manner.

The dataset we used in our research is the “Migration Map” section of “Tencent Location Big Data” [38], with the time interval set to a day and the accuracy able to be traced back to the individual level. The website counts the number of changes in the location of the smart terminal within a certain time interval to filter, summarize, and count the data. In consideration of user privacy, the website only provides the total amount of population inflows and outflows in a day, with the city as the basic unit (the intensity of inflows, source, and outflows, limited to the destination of a single city on a certain day). The website provides a free Application Programming Interface (API) for researchers and programmers, allowing the above data to be obtained and used in scientific research. We used the API with Python programming language to obtain population mobility data during the Spring Festival of 2019 and store it in the SQL database.

The population mobility data we obtained contains the following content (after excluding the user's private information) and were added to a city as the basic units instead of the individual by Tencent company: the source city name and its coordinates, the target city name and its coordinates, time, mobility intensity, and mobility type, which is consistent with the content displayed on the website. After manual filtering and sorting, it contains a total of 40,591 pieces of information, each of which covers eight aspects (source city name and its coordinates, target city name and its coordinates, time, and mobility intensity). Based on this, we constructed a data table of 40,591 * 8, as shown in Table 2-2. Other variables, such as total population, average wage, gross regional product (GRP), urbanization rate, unemployment rate, and other socioeconomic indicators, were obtained from the 2018 urban statistical yearbook on the website of the National Bureau of Statistics. In this chapter, Geographic Information System (GIS), GeoDa, Gephi and GWR4 were used to process, analyze, and calculate the data.

It should also be noted that the error and representativeness of the dataset cannot be ignored. Despite the huge amount of travel data obtained through location services, it is undeniable that there are still some groups who do not use any apps through smartphones, so their travel trajectories cannot be collected. It is predictable, this dataset is representative of specific regions and groups, such as developed regions and young middle-aged groups, which coincides with our research objects. Nonetheless, considering that this dataset provides a larger, more dynamic, and more efficient record of population mobility, combined with fine temporal and spatial resolution, it has proven to be reliable in related research [39-42].

In this study, we considered 290 prefecture-level and above administrative units in China as the research focus, including four municipalities, two special administrative regions, 15 sub-provincial cities and 269 prefecture-level cities. Due to the lack of data, some prefecture-level cities in Hainan province, Taiwan, and some ethnic minority autonomous prefectures in western China were not included in the study. Fig. 4-1 shows the research areas in this chapter.

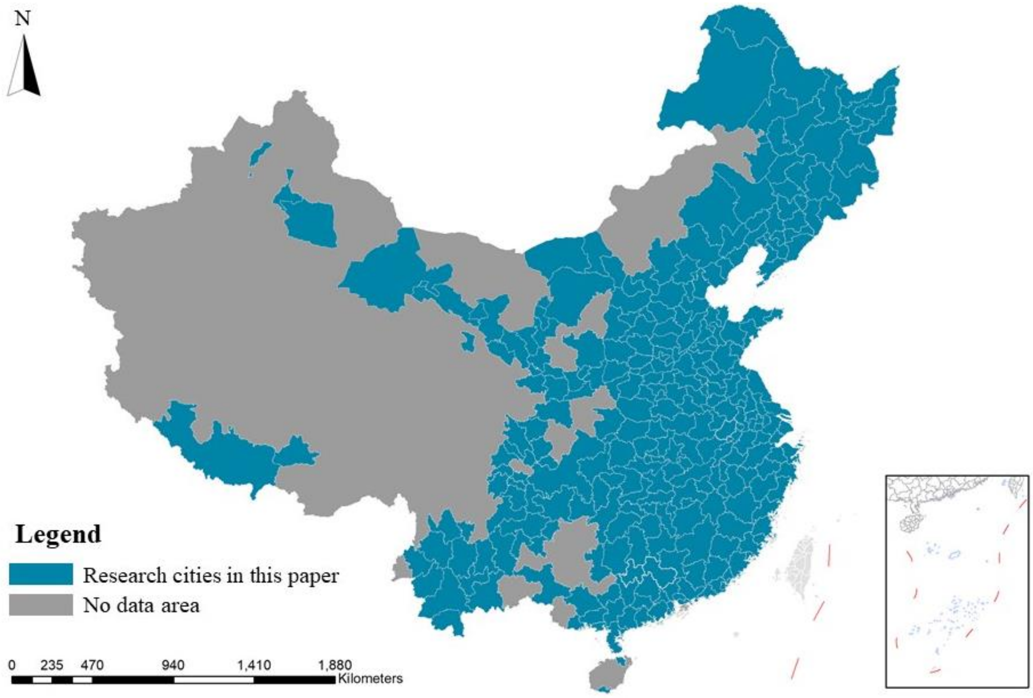


Fig. 4-1 Research areas in this chapter.

4.3 Research methods, variables and model selection.

4.3.1 Research methods

In this study, we first used the social network analysis method to analyze the characteristics of the population mobility network. Then, spatial autocorrelation analysis was used to validate the spatial dependence of population mobility, and the ordinary least squares (OLS) method and correlation test were employed to identify correlated factors of population mobility. After that, three types of regressions analysis, including ordinary least squares (OLS), geographically weighted regression (GWR) and semiparametric geographically weighted regression (SGWR), were conducted to reveal the correlated factors of population mobility. Fig. 2 gives a flowchart of this chapter.

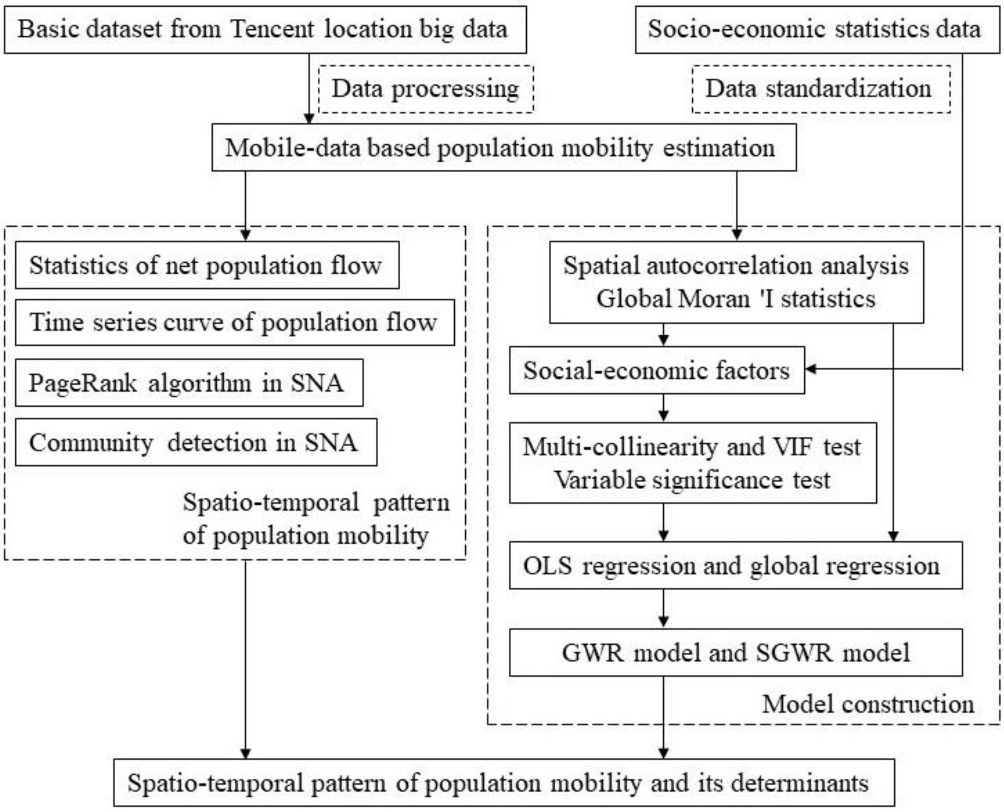


Fig. 4-2 Flow chart of this research.

For detailed introduction of social network analysis and geographically weighted regression model, please refer to chapter 2. Here, we will only give an introduction to the core problems of the GWR model. For the GWR model, the spatial weight matrix is critical. The selection of the spatial weight function has a great influence on the parameter estimation of the geographically weighted regression model [43]. In this chapter, we used an adaptive bi-square kernel to calculate the weight matrix. We used adaptive bi-square kernels instead of fixed kernels based on two considerations. First of all, the regression points (the center of each city) appear to be randomly distributed in the study area, and the adaptive kernel makes the dataset large enough for each local regression [44]. Secondly, the adaptive bi- square kernel can reduce the bandwidth in the data-intensive place and expand the bandwidth in the scattered place of the dataset and have a clear-cut range when the kernel weight is not 0. It has been widely used in studies taking the city as the unit [45, 46]. Meanwhile, the accuracy of the GWR model was greatly affected by the bandwidth of the weight function. Akaike information criterion (AICc) and cross-validation (CV) are two methods commonly used to determine the bandwidth. Compared with the latter, the former can quickly and efficiently resolve differences in the degrees of freedom in various models [47]. Therefore, the AICc was selected to determine the appropriate bandwidth when constructing the GWR model.

4.3.2 Mobile-data-based population mobility variation (MBPMV)

Due to data acquisition, we only get 14 days of Tencent location data from 290 cities, but considering the sample size and data accuracy, this dataset can be used as the basic data for in-depth research. The first day of the dataset was January 29th, and the last day of the dataset was February 11th. We divided the 14 days into two phases, before the Spring Festival and after the Spring Festival. Therefore, the net population mobility in all cities during the two time periods is considered to be representative of the population distribution during the Spring Festival. Through changes in time series and differences in population mobility between the two time periods, we found differences in the distribution of human activity.

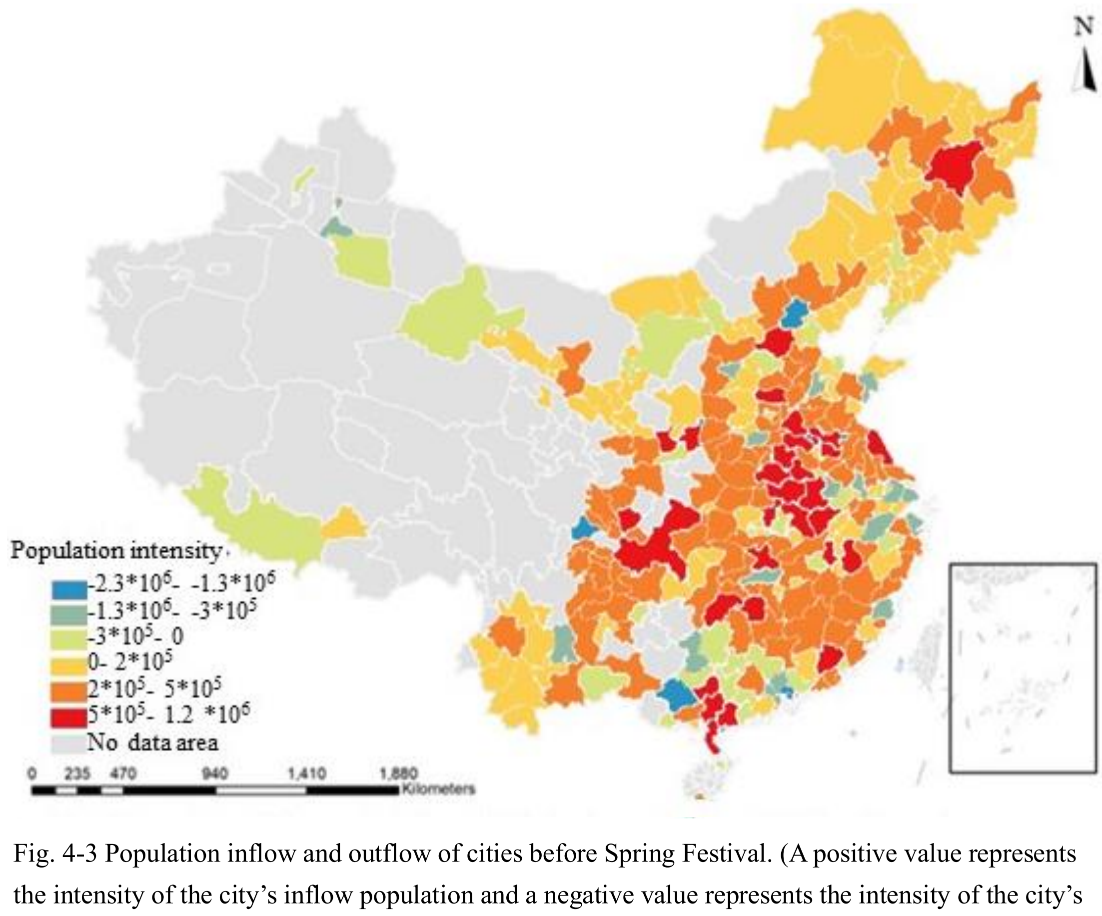


Fig. 4-3 Population inflow and outflow of cities before Spring Festival. (A positive value represents the intensity of the city’s inflow population and a negative value represents the intensity of the city’s outflow population.)

Fig. 4-3 and 4.4 show the inflow and outflow statistics of all cities before and after the Spring Festival, respectively. We selected cities with the most obvious population inflow and outflow in the two time periods, namely Beijing, Chongqing, Shenzhen, and Hengyang, and plotted their main population flow direction and intensity, as shown in chapter (Fig. 2-6). Combining Fig. 4-3 and 4-4, it can be found that the cities with large population mobility are all located in the four major urban agglomerations in China, and the central region experienced a significant population inflow before the Spring Festival and a large population outflow after the Spring Festival. This pattern of population flow reveals the differences in China’s regional development, so it is particularly important to explore the causes of this phenomenon, which may be geopolitical, economic, social, and geographical, etc. Meanwhile, the global spatial autocorrelation analysis of the population mobility changes in all cities after the Spring Festival shows that the Moran index is 0.702, the Z score is 50.128 and the P value is 0.01, indicating that the probability of randomly generating the above spatial distribution pattern is less than 5%. Based on this, a SGWR model can be constructed to explore the socio-economic factors related to the formation of the above population mobility pattern.

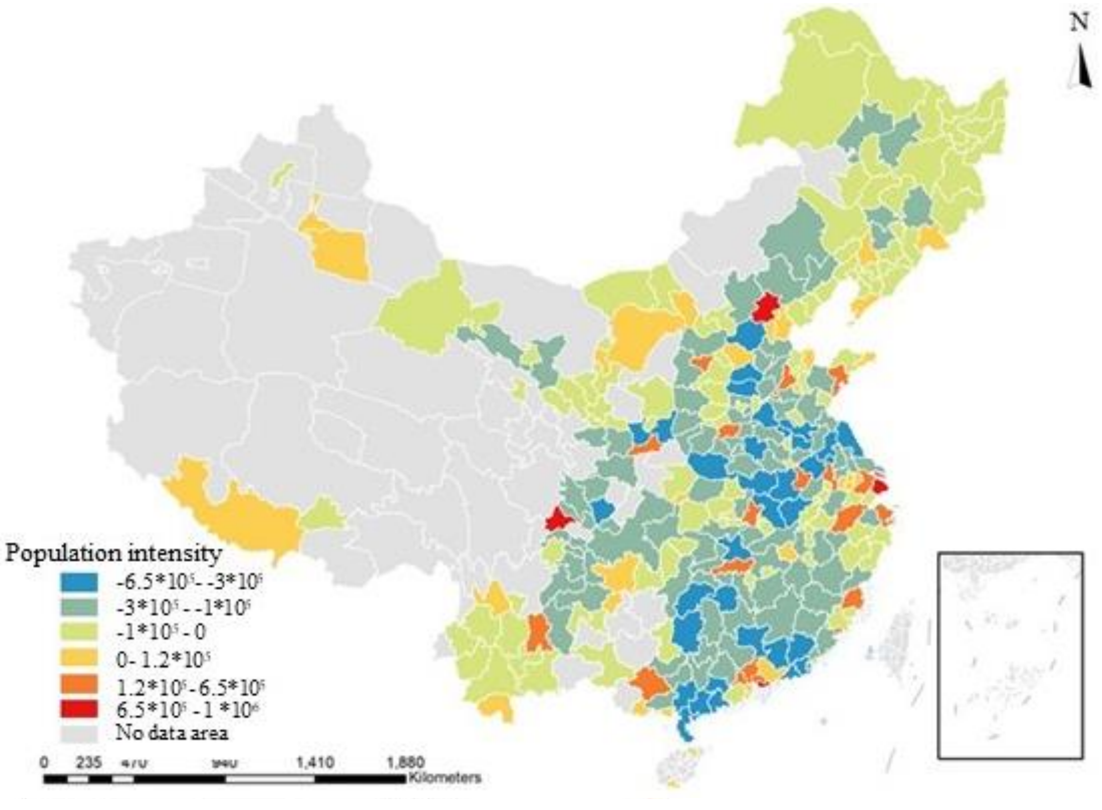


Fig. 4-4 Population inflow and outflow of cities after Spring Festival. (A positive value represents the intensity of the city ’s inflow population and a negative value represents the intensity of the city ’s outflow population.)

4.3.3 Independent variables selection and model construction

In order to analyze the determinants of population mobility patterns, we used three steps to determine the independent variables in GWR model: (1) Select for socio-economic factors related to urban development. The wage level of employees is a very important factor, because the difference in expected benefits is the main force driving population mobility [48]. Secondly, GRP and average gross region product are direct reflections of the economic development of a city, which may affect population mobility [49]. At the same time, depending on different types of work, the three industries can also affect population mobility. Finally, considering that labor-intensive industries can absorb large amounts of labor, we consider foreign capital as a candidate variable [50]. In addition, we added several candidate variables related to urban development and social economy, such as urban total population, urbanization rate, and urban worker unemployment rate [51]. (2) Exclude multicollinearity between variables. We performed an OLS regression to detect multicollinearity between the variables. After all variables are normalized to fit the normal distribution, the variance inflation factor (VIF) of each independent variable was calculated, and then the independent variables with VIF > 7.5 were eliminated from the final model. In this process, the VIF values of the nine independent variables we selected were all less than 7.5, indicating that there was no multicollinearity among the variables (Table 1). (3) Perform a correlation analysis of variables, excluding variables that are not related to population mobility at a 95% confidence level. In this process, the UER was eliminated. The results show that there were no redundant and uncorrelated problems in the remaining variables. Therefore, after the above three processes, TP, AW, GRP, Avg_GRP, UR, FC, VAPI, and VASTI were used for GWR model analysis. Table 2 shows the details of the above variables.

Table 4-1. Variance inflation (VIF) value and correlation coefficient of all independent variables.

	TP	**AW**	**GRP**	**Avg_GRP**	**UR**	**UER**	**FC**	**VAPI**	**VASTI**
VIF	1.236	3.456	6.596	3.408	2.815	1.247	2.641	3.302	2.281
Coefficient	0.163**	0.577**	0.462**	0.405**	0.572**	–0.590	0.582**	–0.503**	0.538**
Sig.	0.003	0.000	0.000	0.000	0.000	0.361	0.000	0.000	0.000

** represents that the variable is significant at the 0.05 level.

After OLS regression and correlation test, a total of eight variables were selected into the GWR model. In this paper, the significance ($p < 0.05$) of all variables was defined as the pseudo t (Est/SE) > 1.96 or < –1.96 [47, 52]. Considering that local models can improve accuracy, the SGWR model was further used to explore the spatial stationarity and non-stationarity of parameters affecting population mobility. An iterative process was used to determine whether the parameters are global or local variables. The most suitable model is judged based on AICc and the model with the smallest AICc value will be selected as the best [53, 54].

Table 4-2. Details of the dependent and independent variables.

Class	**Variable**	**Notation**	**Explanation**
Dependent Variable	MBPM variation	MBPMV	net inflow population of all the cities after the Spring Festival
Independent Variable	Total population	TP	total population at year end (1,000,000 persons)
	Average wage	AW	average wage of employees on duty (yuan/person)

Gross region product	GRP	annual gross regional product (100,000,000 yuan)
Average gross region product	Avg_GRP	annual gross regional product per capita (10,000 yuan/person)
Urbanization rate	UR	proportion of urban population to total population (%)
Unemployment rate	UER	proportion of unemployed population to working population (%)
Foreign capital	FC	actual utilization of foreign investment (100,000,000 yuan)
Value-added of primary industry	VAPI	annual value-added of primary industry (10,000 yuan)
Value-added of secondary and tertiary industry	VASTI	annual value-added of secondary and tertiary industry (10,000 yuan)

4.4 Results

4.4.1 Spatiotemporal patterns of population mobility

Fig. 4-3 and 4-4 show the inflow and outflow statistics of all cities before and after Spring Festival, respectively. We see that: (1) There are significant differences in population mobility in different cities in the two time periods. The population flow shows a high consistency with the city level—that is, the higher the city's development level, the greater its population flow during the Spring Festival, such as the Beijing, Shanghai, Guangzhou, Shenzhen, and Chengdu. (2) The inflow and outflow of population cities are also different in the two time periods. Before the Spring Festival, there was an obvious population inflow in the central and western regions, and the core cities in the four major urban agglomerations had a relatively obvious population outflow. After the Spring Festival, this phenomenon will reverse. This might be because the regional core cities attract a large number of migrant laborers. Before the Spring Festival, the migrant laborers return home to be with their families. This is the commonly mentioned "returning flow." After the Spring Festival, they return to the original work-place to continue their job, which will lead to the so-called "migrant flow", resulting in a surge of migrant laborers to the work place.

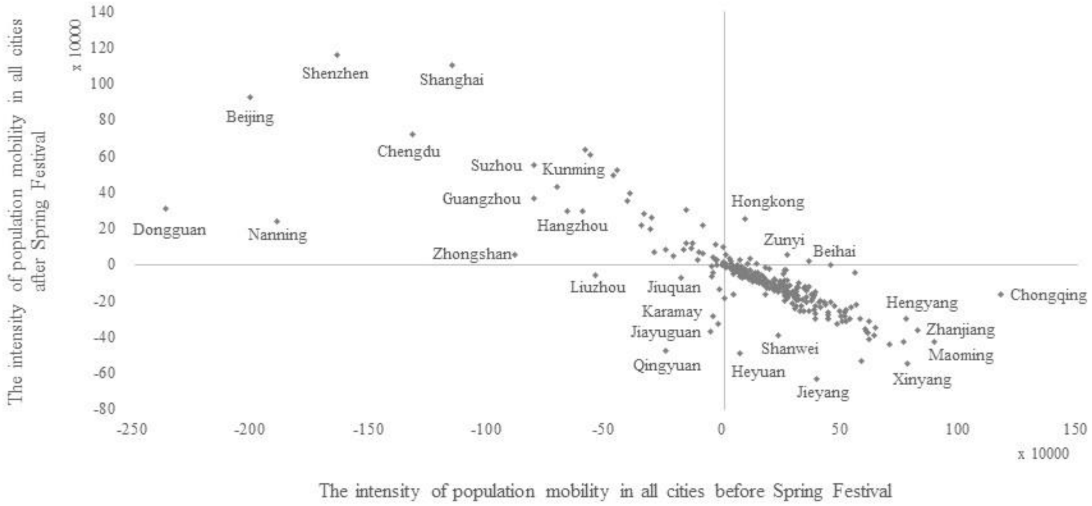


Fig. 4-5 The bivariate relationship between the intensity of population mobility in the two time periods of all the cities.

Based on the statistics of population inflows and outflows in the two periods (Fig. 4-5), we divided all cities into four categories: continuous population inflow (II), continuous population outflow (OO), population inflows then outflow (IO), and population outflows then inflow (OI). Table 3 lists the four different types of cities. We can see that first-tier cities and provincial capital cities located in southeastern China belong to the OI type; second and third-tier cities located in central and western China and small cities around the regional core cities belong to the IO type. Most of the cities with continuous population outflows are located in northwestern China, which has low population attractiveness for economic, environmental, and geographical reasons. The problem of irreversible population loss should attract the attention of the relevant city managers. The same situation also occurred in the Pearl River Delta urban agglomeration. Guangzhou and Shenzhen absorbed a large number of human and material resources from the surrounding areas, resulting in a continuous outflow of the population from the surrounding small cities, which to some extent destroyed the sustainable development of the region. With the improvement of people's living standard, traveling for the New Year has become common. Therefore, a tourism-oriented city such as Sanya, Zunyi, Lijiang, or Beihai can continue to attract visitors to some extent during the Spring Festival.

Table 4-3. Different city types based on population inflow and outflow statistics.

City type	Number	Cities
IO	223	Chongqing, Wenzhou, Xining, Lanzhou, Taizhou, Luoyang, Yangzhou, Xuchang, Shaoxing, and 214 other cities
OI	42	Beijing, Shanghai, Guangzhou, Xiamen, Chengdu, Zhengzhou, Suzhou, Changsha, Jinan, Kunming, Hefei, and 31 other cities
II	12	Hong Kong, Macau, Sanya, Xishuangbanna, Lijiang, Changhzhou, Zunyi, Weihai, Shennong, Langfang, Baishan, and Beihai
OO	13	Karamay, Jiayuguan, Jiuquan, Yangjiang, Qiangjiang, Yunfu, Liuzhou, Jinhua, Haikou, Wuzhou, Chaozhou, Yangjiang, Qingyuan, and Tongchuan

IO represents population inflow the before Spring Festival and outflow after the Spring Festival. OI represents population outflow before the Spring Festival and inflow after the Spring Festival. OO represents continuous population outflow during the Spring Festival; II represents continuous population inflow during the Spring Festival.

Fig. 4-6 is the grading map of the of population flow during the Spring Festival, from which we can clearly see the spatial pattern. First, unlike the diamond-shaped structure formed by the population mobility during the National Day golden week [55], the population flow during the Spring Festival presents a spatial pattern of "two east-west main axes and three north-south main axes." The "two east-west main axes" are Shanghai–Nanjing–Chengdu and Shanghai–Wuhan–Chongqing, and the "three north-south main axes" are Shenzhen–Chengdu, Shenzhen–Wuhan and Guangzhou–Shanghai, all located in the four major urban agglomerations in China. At the same time, we note that, although Beijing is not prominent in this structure, its coverage covers most areas of China and it is also a distributing center for population. The Shandong peninsula is not obvious in this structure, which is badly out of line with its position in the national development strategy. Second, compared with the population flow during the National Day golden week, the population flow boundary of the major cities during the Spring Festival is relatively obvious. Large cities have a typical spatial orientation, while medium-sized cities show strong spatial proximity.

For a further understanding of the spatiotemporal pattern characteristics of population flow, we first established a directed weighted matrix of population inflow and outflow between cities, then explored it by using the PageRank algorithm and "community" detection test in SNA. The PageRank algorithm

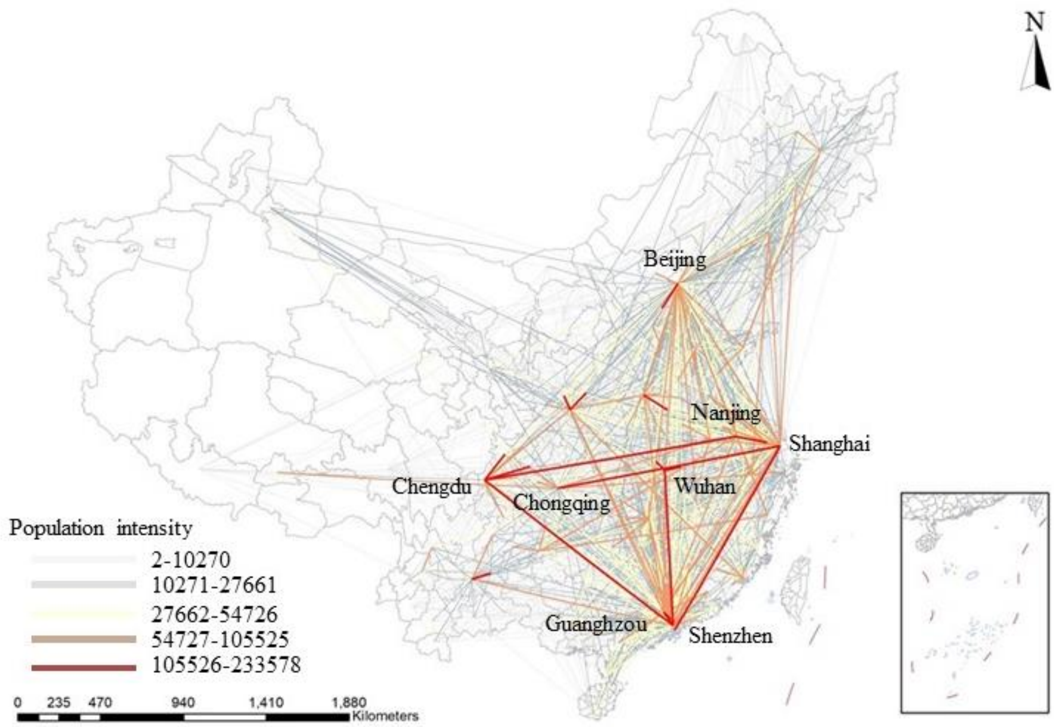


Fig. 4-6. Grading map of the net flow of population during the Spring Festival.

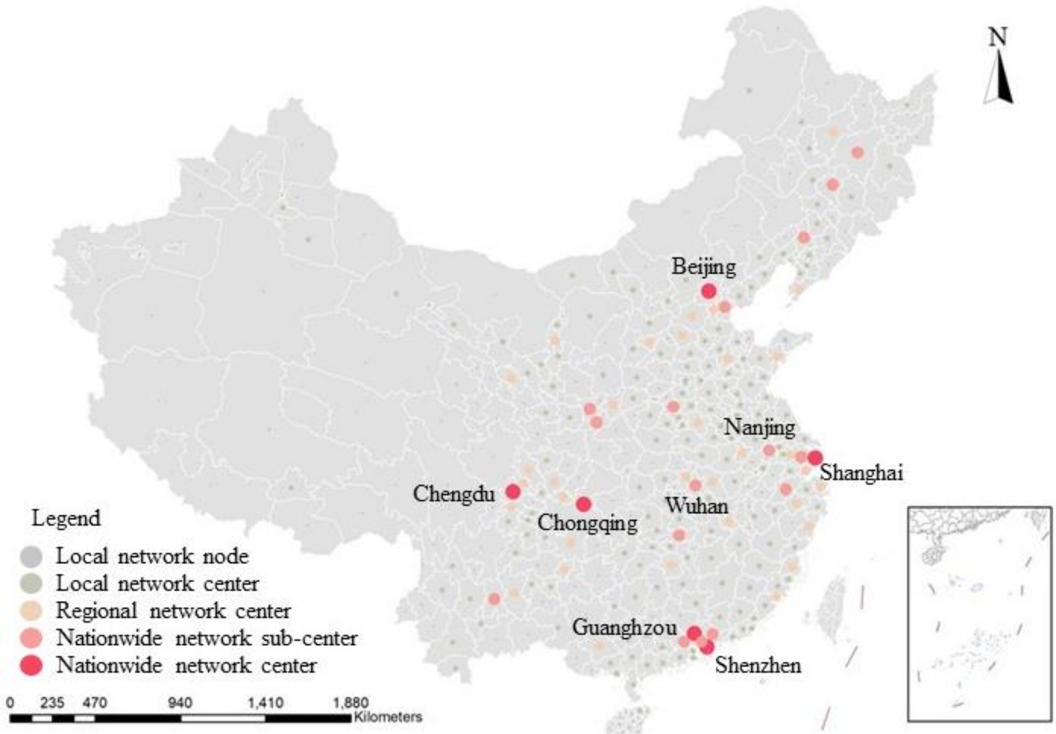


Fig. 4-7. hierarchical map of cities in population mobility network.

was used to rank the importance of cities in the population mobility network, and the hierarchical structure of population flow was obtained. Fig. 4-7 shows a hierarchical map of all cities in the population mobility network, which is classified according to the PageRank value by the natural break classification (NBC); the results are summarized in Table 4-4. We find that there are six cities in the nationwide network center, namely Beijing, Shanghai, Chongqing, Guangzhou, Shenzhen, and Chengdu, all located in the four major urban agglomerations of China, which is similar to the results in Fig. 4-5. The nationwide network subcenter consists of 16 cities, which are either sub-provincial cities, provincial capitals, or developed cities in southeast coastal areas. To a certain extent, the above cities have a clear connection to population mobility during the Spring Festival. Compared with the cities in the southeast coastal areas, the cities in the central and western regions are mostly regional network centers or local network centers, which are not prominent in the whole population mobility network, indicating that the above areas show extremely weak attraction or radiation force in both the population inflow and outflow. From this we can find obvious differences between regions.

Table 4-4. Summary of the city's hierarchy in population mobility network.

Level (PageRank value of network)	**Cities**
Nationwide network center	Beijing, Shanghai, Chongqing, Guangzhou, Shenzhen, and Chengdu
Nationwide network subcenter	Xi'an, Hangzhou, Wuhan, Changsha, Harbin, Zhengzhou, Nanjing, Suzhou, Tianjin, Kunming, Foshan, Huizhou, Shenyang, Changchun, Dongguan, and Xianyang
Regional network center	Baoding, Langfang, Guiyang, Xiamen, Jinan, Ningbo, Nanchong, Sanya, Nanning, Qingdao, Guang'an, Wenzhou, Shijiazhuang, Deyang, Weinan, Hefei, Mianyang, Hongkong, Zunyi, Fuzhou, Meishan, Huanggang and 14 other cities
Local network center	Yueyang, Dali, Shaoxing, Xuzhou, Baoji, Zhaoqing, Yancheng, Shaoguan, and 158 other cities
Local network node	The remaining 66 cities

Through the network analysis method to calculate the matrix of directed weighted population mobility, we find that the clustering coefficient of the population mobility network in the Spring Festival is 0.375, and the average path length is 2.792, which is much higher than the random network composed of 290 nodes (the clustering coefficient is 0.112, and the average path length is 2.075), indicating that the population mobility network during the Spring Festival conforms to the scale-free network characteristics and presents a typical "small world" network structure, which is different from Li's results at the provincial scale [39]. With the help of the "community" detection test, we further revealed the relationship between the cities hidden in the population mobility network. Nodes belonging to the same community tend to be more closely linked, indicating that cities within the same community have more frequent population mobility than other cities. Figure 8 gives the distribution map of network community structure and Table 4-5 summarizes the community structure of all cities.

Based on the analysis of the population mobility network during the Spring Festival, 11 different community structures were identified (Fig. 4-8 and Table 4-5). According to the spatial composition of the community, we divided the 11 communities into three categories: the first is the cross-regional community, such as the community composed of Shanghai, Jiangsu, Zhejiang, Chongqing, and Jilin; the second is the neighborhood community, such as the community composed of Shanxi, Shannxi, Ningxia, and Gansu; the third is independent provinces, such as the community composed of all cities in Shandong province. We found that the second and third community structures accounted for a large proportion of the 11 communities, indicating that large-scale population mobility is still affected by the geographical and geospatial environment. However, like the first community structure, the spatial span was large and distributed across several independent spaces, so it can be seen that, with the improvement of the transportation infrastructure and economic level of the target city, large-scale, cross-regional, and high-density population mobility will become a future development trend, and the space-time distance in the traditional sense will be severely compressed. This reflects the special structure of the population mobility network during the Spring Festival, but we still need to obtain longer time series data for analysis a more general analysis.

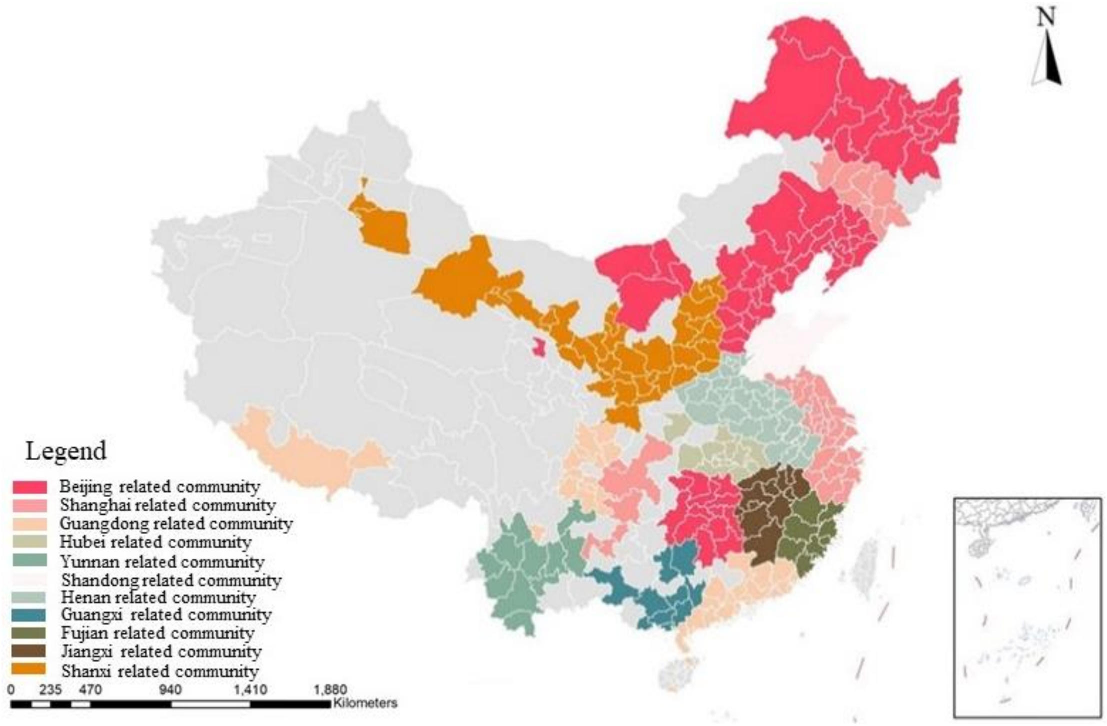


Fig. 4-8 Distribution map of network community structure.

Table 4-5. Summary of the city's community structure.

ID	Major covering provinces	Key cities included	Number of cities
0	Beijing, Tianjin, Hebei, Heilongjiang, Liaoning, Hunan	Beijing, Tianjin, Shijiazhuang, Harbin, Shenyang, Changsha	62
1	Shanghai, Jiangsu, Zhejiang, Chongqing, Jilin	Shanghai, Nanjing, Hangzhou, Chongqing, Changchun	44
2	Tibet, Sichuan, Guangdong	Lhasa, Chengdu, Shenzhen, Guangzhou	38
3	Hubei	Wuhan, Xiangyang	16
4	Yunnan	Kunming, Dali, Qujing	13

5	Shandong	Jinan, Qingdao, Yantai, Weifang	17
6	Henan, Anhui	Zhengzhou, Kaifeng, Luoyang, Hefei	33
7	Guangxi	Nanjing, Liuzhou, Guilin, Wuzhou	10
8	Fujian	Fuzhou, Xiamen, Zhangzhou, Quanzhou	9
9	Jiangxi	Nanchang, Jiujiang, Shangrao, Fuzhou	11
10	Shanxi, Shannxi, Ningxia, Gansu	Xi'an, Xianyang, Urumqi	37

4.4.2 Semiparametric geographically weighted regression (SGWR) model results

Table 4-6 summarizes the basic parameters of the OLS, GWR, and SGWR model outputs. We can see that the constructed SGWR model has made significant improvements over the normal regression model and GWR model. Compared with the traditional regression model, the SGWR model has smaller AICc (472.83) and larger adjusted R2 (0.751), which indicates better overall performance. Also, the F value (2.97) is much higher than the standard value (1.26), which means the null hypothesis that the SGWR model does not improve the traditional regression model can be rejected at the 95% confidence level.

Table 4-6. Summary of OLS, GWR, and SGWR model outputs.

	OLS	**GWR**	**SGWR**
AICc	534.54	480.49	472.83
Adjusted R-squared	0.633	0.748	0.751
Bandwidth	—	98.72	89.70
Residual squares	101.40	53.60	50.01

ANOVA

Source	**SS**	**DF**	**MS**	**F**	**F Criterion**
Global Residuals	101.40	279			
GWR Improvements	47.80	64.37	0.74		
GWR Residuals	53.60	214.63	0.25	2.97[a]	1.26

[a]Statistically significant at a confidence level of 95%.

Tables 4-7 and 4-8 illustrate the statistics of the SGWR model and global regression model output.

The results showed that AW, UR, FC, VAPI, and VASTI were significant at the 95% confidence level, while TP, GRP, and Avg_GRP were not significant at the 0.05 confidence level. Among them, VASTI and MBPMV have the strongest positive correlation; VAPI and MBPMV have the strongest negative correlation; and FC, UR, and AW also have a strong positive correlation with MBPMV. TP, GRP, and Avg_GRP are not significantly correlated with MBPMV. Meanwhile, UR was finally selected as a global parameter after an iterative process in GWR4.

Table 4-7. Summary of the global regression model outputs.

Variable	**Estimate**	**Standard Error**	**Pseudo *t***	**t-test**	**Spatial stationarity**
Intercept	0.000	0.103	0.000	-	-
TP	0.008	0.038	0.217	*P>0.05*	-
GRP	0.021	0.063	0.339	*P>0.05*	-
Avg_GRP	–0.093	0.063	–1.472	*P>0.05*	-
UR	0.138	0.060	2.301	*P<0.05*	global
AW	0.140	0.058	2.422	*P<0.05*	local
FC	0.280	0.056	4.967	*P<0.05*	local
VAPI	–0.814	0.070	–11.505	*P<0.05*	local
VASTI	1.127	0.106	10.622	*P<0.05*	local

Table 4-8. Summary of the SGWR model estimation coefficients.

Variable	**Mean**	**Std.**	**Min**	**Lwr Quartile**	**Median**	**Upr Quartile**	**Max**
Intercept	–0.320	1.366	–5.810	–0.215	–0.093	0.004	2.199
TP	–0.598	0.431	–1.339	–0.906	–0.643	–0.146	0.036
AW	2.120	1.257	0.099	1.010	1.935	3.237	4.645
GRP	0.621	1.025	–1.549	–0.140	0.586	1.406	3.265
Avg_GRP	–0.034	0.126	–0.251	–0.130	–0.039	0.044	0.296
FC	1.65	2.451	–3.98	–0.20	2.10	3.58	5.88
VAPI	–2.649	1.758	–6.133	–4.011	–2.451	–1.309	0.339
VASTI	1.02	2.312	–3.233	–0.64	0.17	2.315	5.971

Figures 4-9 to 4-13 visualize the spatial variation and estimation coefficients of all explanatory variables. Figure 9 shows that there is a large spatial difference in the value of local R^2, which indicates that, with the change in urban spatial location, explanatory variables have different interpretation forces on dependent variables, further reflecting the spatial nonstationarity between variables. In addition, the standard residual of the model was analyzed and the model presented a random distribution pattern in space, indicating that the constructed SGWR model had better performance.

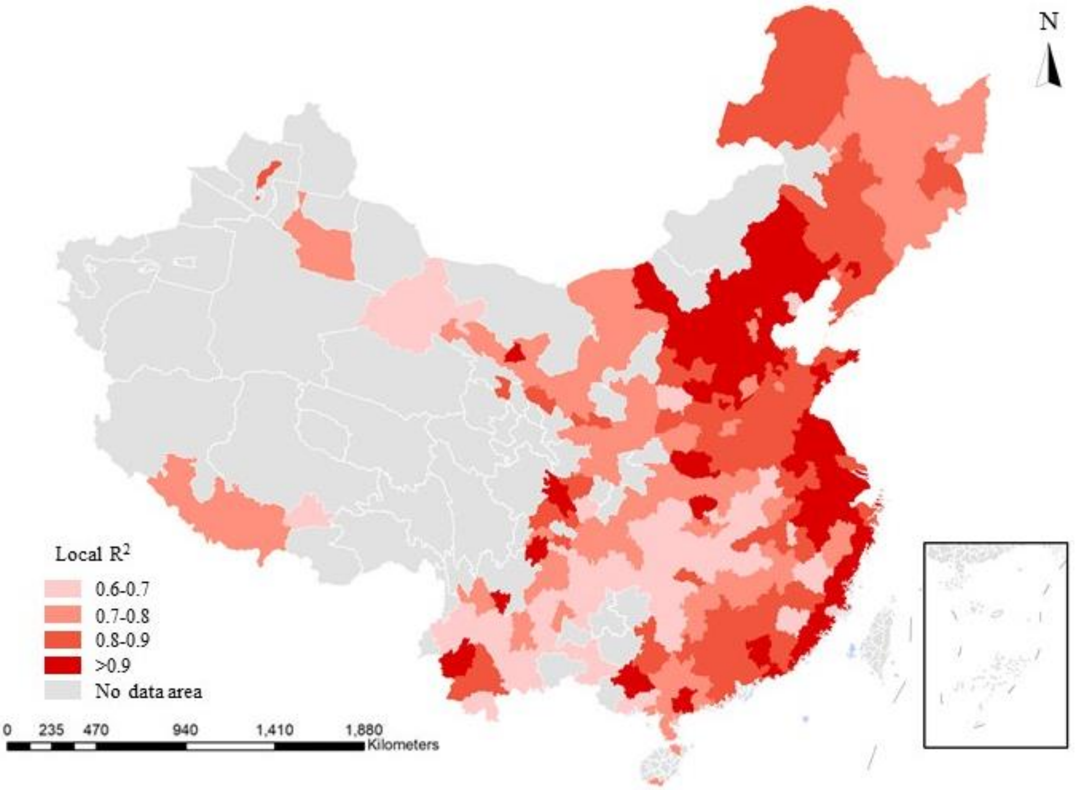


Fig. 4-9 Local R^2 based on SGWR model.

According to the statistical results of the model, the added value of the secondary and tertiary industries, the wage level of employees, the urbanization rate, and foreign capital are positively correlated to the population mobility. The added value of the primary industry is negatively correlated with the population mobility. There has no significant correlation between the total population, unemployment rate, GRP, and population mobility. In addition to the urbanization rate, other variables have different effects on different regions. These results are basically consistent with reality, as explained in the following.

The strongest positive correlation between VASTI and MBPMV indicates that the added value of secondary and tertiary industries has a significant effect on population mobility. This is because our research focuses on the Spring Festival, during which the work flow is in an absolute position in the population mobility, representing the transfer of labor. Therefore, with the rapid development of the secondary and tertiary industries, the city will provide a large number of jobs, able to absorb the labor force in the surroundings and even farther afield. The population of underdeveloped areas will shift to developed areas, and the population of poor areas will shift to less developed areas. This progressive relationship affects population mobility in all areas.

The average wage of employees also has a positive correlation with MBPMV. This is because, as neoclassical theorists explain, the income level of the intended destination is the main driver force of the migration process. Therefore, when other costs are constant and incomes increase, more laborers will choose higher-paying areas for employment, which is similar to the impact of the added value of secondary and tertiary industries on MBPMV.

Total foreign capital also has a positive impact on population mobility. In most cases, overseas investment aims at the development of secondary and tertiary industries in the city, combined with the construction of labor–intensive enterprises, directly creating a large number of positions in the city, so this economic factor also increases population mobility.

Urbanization is also positively correlated with MBPMV. With the increase in the urbanization level, on the one hand, industrial industry can be effectively developed and more employment opportunities will be created through the intensive use of infrastructure. On the other hand, this accelerates residents' socialization and promotes developments in the service industry, which will also create a large number of employment opportunities.

There is a significant negative correlation between the added-value of the primary industry and MBPMV, which indicates that, with the increase in primary industry, the population outflow will be intensified. This is determined by the nature of the work in primary industry. In China, agriculture, forestry, animal husbandry, and fisheries are classified as primary industries. In the context of mechanization, they cannot provide a large amount of labor or even absorb local surplus labor, resulting in population movement elsewhere.

In the above paragraph, we explained the explanatory variables related to MBPMV. Considering that the model takes into account the non-stationarity of space; we will focus on explaining the variation of variables in space as follows.

From Fig. 4-10, we see that the development of the secondary and tertiary industries in the eastern coastal areas has a positive correlation with population mobility, which indicates that, if investment in the secondary and tertiary industries is increased in Jiangsu, Zhejiang, and Shanghai, it could attract more of the floating population. Meanwhile, there is a weak negative correlation between the Beijing–Tianjin–Hebei region and the Pearl River Delta region, which indicates that these two regions will not attract people through the development of secondary and tertiary industries. Some cities in the central and western regions have relatively obvious negative coefficients, especially those in Hunan, Hubei, Yunnan, Guangxi and Henan, all of which are major labor-outputting provinces, and the development of secondary and tertiary industries will not be attractive to the population mobility. Sichuan and Chongqing are also major labor-outputting regions, but their population mobility has a positive correlation with the city's secondary and tertiary industries, which means that population outflow could be slowed by increasing the proportion of secondary and tertiary industries.

Fig. 4-11 implies that there are both positive and negative correlations between foreign capital and population mobility. The Beijing–Tianjin–Hebei region, as well as Zhejiang and Fujian, are the most obvious, which means that the increase in foreign capital can not only reduce the local population outflow, but also attract a large population inflow. Considering that the above regions are the most developed regions of the country as well as the places where talent concentrates, the high-tech industries directly funded by foreign capital can further absorb the talent in the surrounding areas, resulting in a large population inflow. Large negative regression coefficients exist in Henan, Anhui,

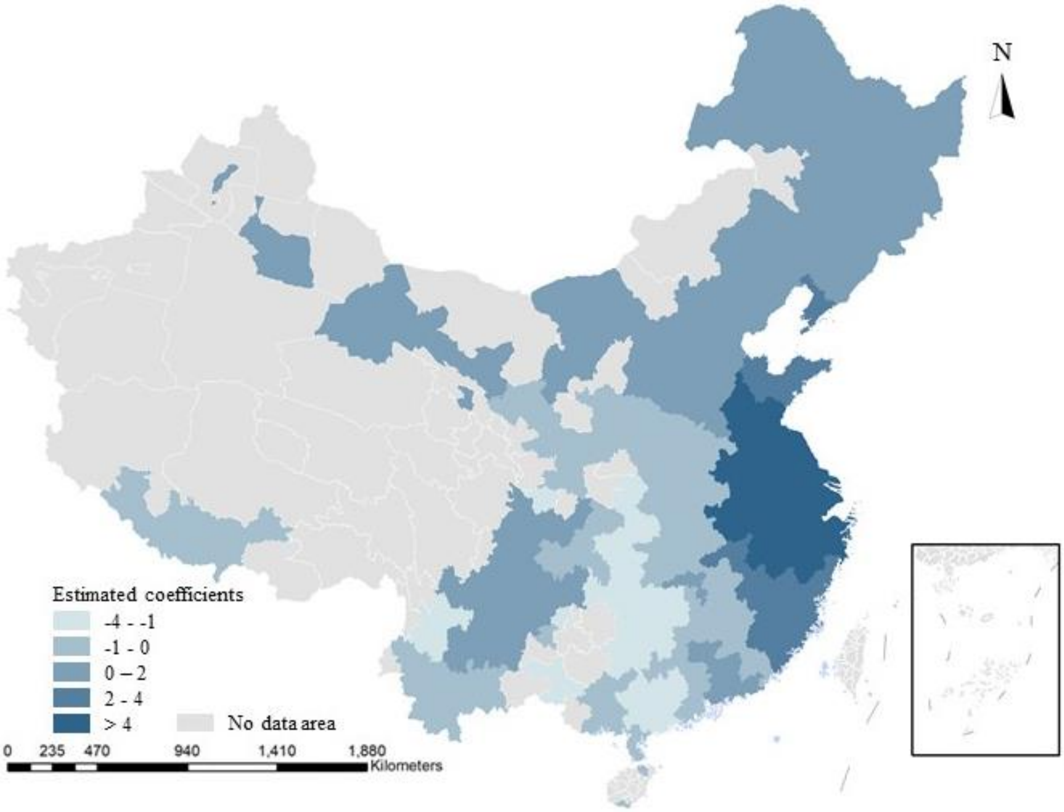


Fig. 4-10. Local $R^2$of value-added secondary and tertiary industry.

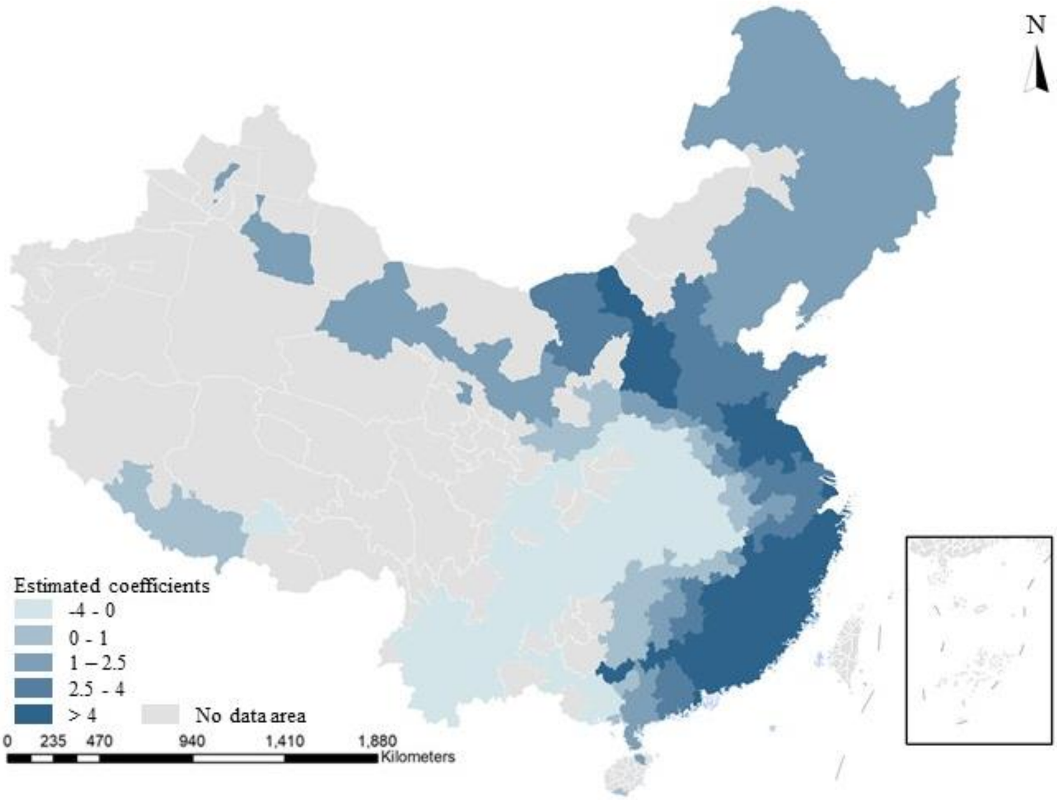


Fig. 4-11. Local R^2 of foreign capital.

Sichuan, and Chongqing, indicating that increasing foreign capital might not a good measure to attract population inflow for provinces with a large labor force output. There are weak positive and negative regression coefficients in the Northeast and Southwest, which may mean that they are already saturated with foreign investment, and an increase will not attract further migration.

From Fig. 4-12, we find that the value added of the primary industry is negatively correlated with the population mobility in all cities studied, which is consistent with a recent study that examined the effects of rising agricultural productivity on migration [56]. The nature of the primary industry means that it can solve the local surplus labor to a certain extent, but it cannot attract external population. The correlation between the central and eastern coastal areas is much higher than that of other regions, which indicates that the above-mentioned regions, especially those of Anhui, Jiangsu and, Zhejiang, should focus on reducing the development of the primary industry in the hopes of attracting external population.

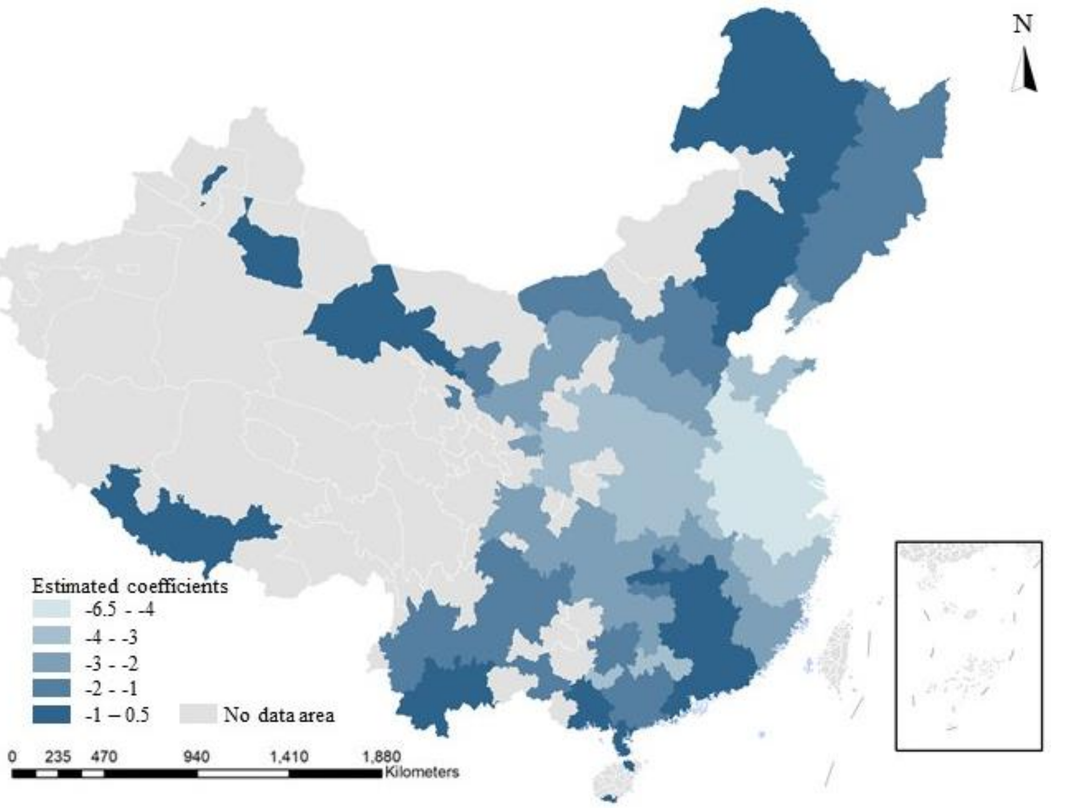


Fig. 4-12. Local R^2 of value-added primary industry.

From Fig. 4-13, we see that the average wage of employees is positively correlated with population mobility in all the cities studied. The coefficient of all cities in southern China is higher than that of the north, which means that the average wage of employees in the southern region, especially in Hunan, Guangdong, and Fujian, is more closely related to population mobility. These regions can attract population inflows by increasing the income of employees. The Beijing–Tianjin–Hebei region and central Shaanxi, Sichuan, Hubei, and other cities have a weak positive correlation coefficient. The former may mean that, even with the further increase in wages, it has not been able to attract a population inflow, while the latter group is mostly labor-outputting cities, perhaps due to the increase in local wages still not reaching the level of developed regions.

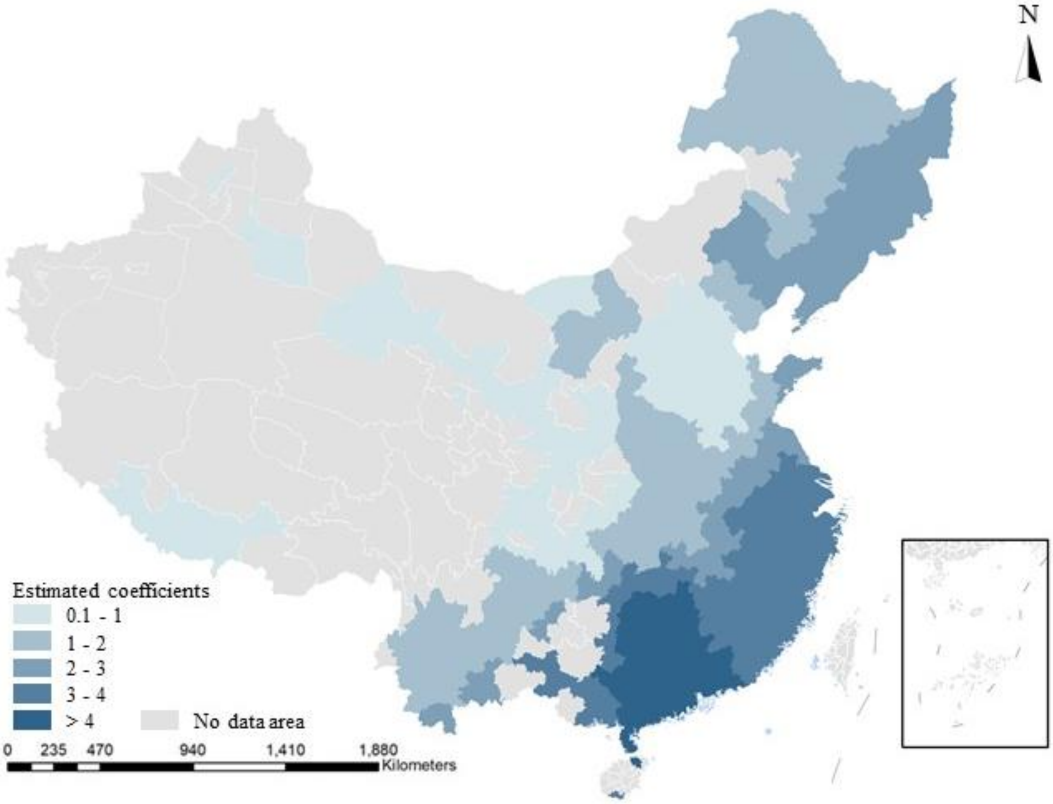


Fig. 4-13. Local R^2 of average wage.

4.5 Summary

Traditional census data cannot reveal the spatial patterns of population mobility and relevant socioeconomic factors within a specific period or even track the people's trajectories due to the slow updating frequency and other shortcomings. Secondly, China's published population distribution statistics often have problems such as low granularity and a poor refinement level, which means daily population movement cannot be described with high spatial and temporal resolution. The spatiotemporal location big data explored the route and direction of population mobility in a relatively continuous time interval, which provided new data support for the study of population distribution and population mobility. Different from the macro model under the long-term evolution rule of statistical data, the research based on travel big data can not only reflect the new characteristics of population inflow and outflow between cities, and describe the agglomeration and diffusion of population flow, but also analyzes the increasingly complex relationships between cities from the perspective of flow through space.

Based on the Tencent application dataset, this study first used the social network analysis method to explore the spatial and temporal distribution patterns and characteristics of population mobility during the Spring Festival, then constructed a SGWR model to reveal the socioeconomic factors related to population mobility. Different from the diamond-shaped structure proposed by Pan and Lai [55], the population mobility network during the Spring Festival presents a typical structure of " two east-west main axes and three north-south main axes." The vertices of the structure are all located in the four major urban agglomerations of China, which reflects the great attraction of the above areas. The social network analysis method not only identified different "community" structures, but also classified all cities hierarchically, reflecting the status of different cities in the population mobility network. The results of the SGWR model show that population mobility is significantly correlated with regional average wage level, urbanization rate, foreign capital, value added of primary industry, and value added of secondary and tertiary industry, which is consistent with findings of Zhong [57] and Li [39].

Using refined datasets with high temporal and spatial resolution, this study explores the structural characteristics of population mobility networks and the heterogeneity of different regions in attracting populations, thus surpassing previous studies. We found that the population tends to shift from low-wage, low-input and low-development level areas to high-wage, high-input and high-development level areas. On the one hand, this supports the early analysis conclusion based on census data—that economic differences between regions are the main force driving population mobility [10, 11]. On the other hand, it validates the neoclassical migration theory that migrants make their decision to move as a response to interregional or rural–urban wage differentials and rational cost–benefit calculations [6, 7]. The imbalance between developments in this region is due to the dual urban–rural development in the planned economy era and the priority development of the eastern region after the reform and opening-up [16, 37]. The imbalance in regional development is directly manifested in the imbalance of economic development, which intensifies the scale of population mobility and ultimately leads to a growing gap between regions. Meanwhile, such large-scale movement is also a huge challenge to the transportation system. A series of social problems such as left-behind children, poor living conditions, and urban diseases are often hidden behind the large-scale and long-term population mobility. Although we are unable to solve these problems, we can propose some positive development and management strategies through the analysis of population mobility. First, using location big data to track population activity trajectories and investigate population distribution is critical to the good operation of urban systems. Secondly, different economic and social development strategies should be implemented according to the development status and conditions to avoid excessive population loss and the “shrinking city” phenomenon. For major labor-outputting regions such as Sichuan and Shandong, the former can try industrial transformation to speed up the development of secondary and tertiary industries, the latter can increase the intensity of attracting foreign investment, and the central region should consider increasing the income level of local employees.

The limitations of this chapter cannot be ignored. First, there will inevitably be problems such as data deviation, data discontinuity, and data loss. As mentioned earlier, Tencent has hundreds of millions of users, but there are still some people who do not use products developed by Tencent, so their travel behaviors will not be recorded. Due to privacy issues, the original dataset does not provide the social attributes of population (gender, age, occupation, and purpose), so we cannot accurately assess the purpose of the population movement (the majority of it is migrant worker flow, but there is still some student and tourism flow). The support of multi-source data and the cross-application of multi-disciplinary fields are the key to studying cities and population in the era of big data. Second, population mobility is restricted and influenced by many complex factors. In this paper, socioeconomic factors were explored with the help of the SGWR model. Recent studies have investigated amenity-led, public policy-led and tourism industry-led population mobility and indicated that the above two factors play an increasingly important role in attracting population mobility [50, 58-60]. In the next stage, the relationships between the quality of life factor, amenity factor and public factor and differences in regional population mobility need to be explored. Third, we only analyzed the 14-day population mobility in this paper due to the confidentiality of the data. It is well known that there is also a peak period of population mobility before and after the Lantern Festival (the 15th day of the 1st lunar month), which was not reflected in this paper. Therefore, in future studies, population mobility data with longer time series should be obtained and analyzed, and the results may be more representative and instructive. This paper combined the social network analysis method and the SGWR model to explore the spatiotemporal patterns and characteristics of population mobility, thereby revealing the socioeconomic factors related to population mobility. The research results can provide data support for urban policy makers and researchers, while also promoting progress in studies on population mobility.

Reference

1. De Haas, H., Migration and development: A theoretical perspective. *International migration review* **2010,** 44, (1), 227-264.
2. Ma, H.; Chen, Z., Patterns of inter-provincial migration in China: Evidence from the sixth population census. *Population Research* **2012,** 36, (6), 87-99.
3. Ravenstein, E. G., The laws of migration. *Journal of the statistical society of London* **1885,** 48, (2), 167-235.
4. Bogue, D. J., The use of place-of-birth and duration-of-residence data for studying internal migration. *Ekistics* **1960,** 9, (52), 97-101.
5. Lewis, W. A., Economic development with unlimited supplies of labour. *The manchester school* **1954,** 22, (2), 139-191.
6. Ranis, G.; Fei, J. C., A theory of economic development. *The american economic review* **1961**, 533-565.
7. Sjaastad, L. A., The costs and returns of human migration. *Journal of political Economy* **1962,** 70, (5, Part 2), 80-93.
8. Knapp, T. A.; Gravest, P. E., On the role of amenities in models of migration and regional development. *Journal of regional science* **1989,** 29, (1), 71-87.
9. Mueser, P. R.; Graves, P. E., Examining the role of economic opportunity and amenities in explaining population redistribution. *Journal of Urban Economics* **1995,** 37, (1), 1-25.
10. Fan, C. C., Modeling interprovincial migration in China, 1985-2000. *Eurasian Geography and Economics* **2005,** 46, (3), 165-184.
11. Liu, Y.; Shen, J., Modelling skilled and less ‐ skilled interregional migrations in China, 2000 – 2005. *Population, Space and Place* **2017,** 23, (4), e2027.
12. Fang, C.; Dewen, W., Migration as marketization: What can we learn from China's 2000 census data? *China Review* **2003**, 73-93.
13. Gu, C.; Cai, J.; Zhang, W.; Ma, Q.; Chan, R.; Li, W.; Shen, D., A study on the patterns of migration in Chinese large and medium cities. *Acta Geogaphica Sinica* **1999**, (3), 66.
14. Wang, G.; Pan, Z.; Lu, Y., China’s inter-provincial migration patterns and influential factors: Evidence from year 2000 and 2010 population census of China. *Chinese Journal of Population Science* **2012,** 5, 2-13.
15. Yu, T., Spatial-temporal features and influential factors of the China urban floating population growth. *Chinese Journal of Population Science* **2012,** 4, 47-58.
16. Shen, J., Increasing internal migration in China from 1985 to 2005: Institutional versus economic drivers. *Habitat International* **2013,** 39, 1-7.
17. Liao, F. H.; Wei, Y. D., Space, scale, and regional inequality in provincial China: A spatial filtering approach. *Applied Geography* **2015,** 61, 94-104.
18. Wei, Y. D.; Ye, X., Beyond convergence: Space, scale, and regional inequality in China. *Tijdschrift voor economische en sociale geografie* **2009,** 100, (1), 59-80.
19. Liu, Y.; Stillwell, J.; Shen, J.; Daras, K., Interprovincial migration, regional development and state policy in China, 1985–2010. *Applied Spatial Analysis and Policy* **2014,** 7, (1), 47-70.
20. ZHONG, F.-n.; Qing, L.; Xiang, J.; Jing, Z., Economic growth, demographic change and rural-urban migration in China. *Journal of Integrative Agriculture* **2013,** 12, (10), 1884-1895.
21. He, J., The regional concentration of China's interprovincial migration flows, 1982–90. *Population and Environment* **2002,** 24, (2), 149-182.

22. Liu, C.; Otsubo, K.; Wang, Q.; Ichinose, T.; Ishimura, S., Spatial and temporal changes of floating population in China between 1990 and 2000. *Chinese Geographical Science* **2007,** 17, (2), 99-109.
23. Morrill, R., Fifty years of population change in the US 1960–2010. *Cities* **2012,** 29, S29-S40.
24. Wang, X.-R.; Hui, E. C.-M.; Sun, J.-X., Population migration, urbanization and housing prices: Evidence from the cities in China. *Habitat International* **2017,** 66, 49-56.
25. Cao, G.; Liu, T., Rising role of inland regions in China's urbanization in the 21st century: The new trend and its explanation. *Acta Geographica Sinica* **2011,** 66, (12), 1631-1643.
26. Ding, J.; Liu, Z.; Cheng, D.; Liu, J.; Zou, J., Areal differentiation of inter-provincial migration in China and characteristics of the flow field. *Acta Geographica Sinica* **2005,** 60, (1), 106-114.
27. Shen, J., Changing patterns and determinants of interprovincial migration in China 1985–2000. *Population, Space and Place* **2012,** 18, (3), 384-402.
28. Wang, Y.; Li, H.; Yu, Z.; Luo, B., Approaches to census mapping: Chinese solution in 2010 rounded census. *Chinese geographical science* **2012,** 22, (3), 356-366.
29. Bo, W.; Feng, Z.; Hao, Z., The dynamic changes of urban space-time activity and activity zoning based on check-in data in Sina web. *Scientia Geographica Sinica* **2015,** 35, (2), 151-160.
30. Qin, X.; Zhen, F.; Zhu, S.; Xi, G., Spatial pattern of catering industry in Nanjing urban area based on the degree of public praise from internet: A case study of dianping. com. *Scientia Geographica Sinica* **2014,** 34, (7), 810-817.
31. Xiao, Q.; Feng, Z.; Lifang, X.; Shoujia, Z., Methods in urban temporal and spatial behavior research in the big data era. *Progress in geography* **2013,** 32, (9), 1352-1361.
32. Long, Y.; Zhang, Y.; Cui, C., Identifying commuting pattern of Beijing using bus smart card data. *Acta Geographica Sinica* **2012,** 67, (10), 1339-1352.
33. Otte, E.; Rousseau, R., Social network analysis: a powerful strategy, also for the information sciences. *Journal of information Science* **2002,** 28, (6), 441-453.
34. Ivaldi, M.; Ferreri, L.; Daolio, F.; Giacobini, M. D. L.; Tomassini, M.; Rainoldi, A. In *We-Sport: from academic spin-off to data-base for complex network analysis; an innovative approach to a new technology*, 3rd National Congress of Italian Society of Movement and Sport Sciences, 2011; 2011; pp 53-53.
35. Hu, X.; Li, H.; Bao, X. In *Urban population mobility patterns in Spring Festival Transportation: Insights from Weibo data*, 2017 International Conference on Service Systems and Service Management, 2017; IEEE: 2017; pp 1-6.
36. Wei, Y.; Song, W.; Xiu, C.; Zhao, Z., The rich-club phenomenon of China's population flow network during the country's spring festival. *Applied Geography* **2018,** 96, 77-85.
37. China Internet Report 2019. https://www.abacusnews.com/china-internet-report (20 March),
38. Tencent Location Big Data. https://heat.qq.com/index.php (accessed on 12 March 2019),
39. Li, J.; Ye, Q.; Deng, X.; Liu, Y.; Liu, Y., Spatial-temporal analysis on Spring Festival travel rush in China based on multisource big data. *Sustainability* **2016,** 8, (11), 1184.
40. Wang, Y.; Dong, L.; Liu, Y.; Huang, Z.; Liu, Y., Migration patterns in China extracted from mobile positioning data. *Habitat International* **2019,** 86, 71-80.
41. Xu, J.; Li, A.; Li, D.; Liu, Y.; Du, Y.; Pei, T.; Ma, T.; Zhou, C., Difference of urban development in China from the perspective of passenger transport around Spring Festival. *Applied Geography* **2017,** 87, 85-96.
42. Zhu, D.; Huang, Z.; Shi, L.; Wu, L.; Liu, Y., Inferring spatial interaction patterns from sequential

snapshots of spatial distributions. *International Journal of Geographical Information Science* **2018,** 32, (4), 783-805.

43. Brunsdon, C.; Fotheringham, A. S.; Charlton, M. E., Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis* **1996,** 28, (4), 281-298.
44. Feuillet, T.; Salze, P.; Charreire, H.; Menai, M.; Enaux, C.; Perchoux, C.; Hess, F.; Kesse-Guyot, E.; Hercberg, S.; Simon, C., Built environment in local relation with walking: Why here and not there? *Journal of Transport & Health* **2016,** 3, (4), 500-512.
45. Jin, C.; Xu, J.; Huang, Z., Spatiotemporal analysis of regional tourism development: A semiparametric Geographically Weighted Regression model approach. *Habitat International* **2019,** 87, 1-10.
46. Lee, K. H.; Heo, J.; Jayaraman, R.; Dawson, S., Proximity to parks and natural areas as an environmental determinant to spatial disparities in obesity prevalence. *Applied Geography* **2019,** 112, 102074.
47. Fotheringham, A. S.; Brunsdon, C.; Charlton, M., *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons: 2003.
48. Liu, T.; Qi, Y.; Cao, G., China's floating population in the 21st century: Uneven landscape, influencing factors, and effects on urbanization. *Acta Geographica Sinica* **2015,** 70, (4), 567-581.
49. Cao, Z.; Zheng, X.; Liu, Y.; Li, Y.; Chen, Y., Exploring the changing patterns of China's migration and its determinants using census data of 2000 and 2010. *Habitat International* **2018,** 82, 72-82.
50. Liu, Y.; Shen, J., Jobs or amenities? Location choices of interprovincial skilled migrants in China, 2000–2005. *Population, Space and Place* **2014,** 20, (7), 592-605.
51. Zhang, K. H.; Shunfeng, S., Rural–urban migration and urbanization in China: Evidence from time-series and cross-section analyses. *China Economic Review* **2003,** 14, (4), 386-400.
52. Nakaya, T.; Fotheringham, S.; Charlton, M.; Brunsdon, C., Semiparametric geographically weighted generalised linear modelling in GWR 4.0. **2009**.
53. Bozdogan, H., Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika* **1987,** 52, (3), 345-370.
54. Burnham, K. P.; Anderson, D. R., Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research* **2004,** 33, (2), 261-304.
55. Pan, J.; Lai, J., Spatial pattern of population mobility among cities in China: Case study of the National Day plus Mid-Autumn Festival based on Tencent migration data. *Cities* **2019,** 94, 55-69.
56. Yang, J.; Wang, H.; Jin, S.; Chen, K.; Riedinger, J.; Peng, C., Migration, local off-farm employment, and agricultural production efficiency: evidence from China. *Journal of Productivity Analysis* **2016,** 45, (3), 247-259.
57. Zhong, C.; Arisona, S. M.; Huang, X.; Batty, M.; Schmitt, G., Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science* **2014,** 28, (11), 2178-2199.
58. Cai, G.; Xu, L.; Gao, W.; Hong, Y.; Ying, X.; Wang, Y.; Qian, F., The Positive Impacts of Exhibition-Driven Tourism on Sustainable Tourism, Economics, and Population: The Case of the Echigo–Tsumari Art Triennale in Japan. *International Journal of Environmental Research and Public Health* **2020,** 17, (5), 1489.

59. Wu, J.; Yu, Z.; Wei, Y. D.; Yang, L., Changing distribution of migrant population and its influencing factors in urban China: Economic transition, public policy, and amenities. *Habitat International* **2019,** 94, 102063.

60. Liu, Y.; Shen, J., Spatial patterns and determinants of skilled internal migration in C hina, 2000–2005. *Papers in Regional Science* **2014,** 93, (4), 749-771.

Chapter 5. Empirical Analysis on the Mechanism of Population Mobility Promoting Urban Development in China

5.1 Introduction

Since the reform and opening up, China's economy has achieved world-renowned achievements. The economic aggregate has reached a new level and has grown into the world's second largest economy after the United States [1]. The per capita GDP has continued to increase, successfully achieving a leap from low-income countries to middle-income countries. The industrial structure is continuously optimized and adjusted, gradually shifting from an industrial-led industrial system to a service-led modern industrial system [2]. The economic structure has gradually shifted from investment-oriented and consumption-assisted to consumption-oriented and investment-assisted. After decades of development and continuous adjustment, China's economy has entered a "new normal" stage. This means that for a period of time in the future, China's economy will shift from the traditional rapid growth to the stable growth stage [3]. The reconstruction of the economic structure and the reshaping of development momentum will stabilize the growth of China's economy.

Under the realistic background that the national economic development as a whole has entered the "new normal", the difference in regional economic development is still the most prominent problem in China's economy and society. The development pattern of strong east and weak west has always existed and is constantly deepening. After the reform and opening up, the eastern and southeast coastal areas, especially the Pearl River Delta, Yangtze River Delta and Beijing-Tianjin-Hebei urban agglomerations have developed ahead of the central and western regions by virtue of their inherent location advantages and policy design. With the acceleration of the process of urbanization and industrialization, China has experienced a process of transition from an agricultural economy to an industrial economy and then to a post-industrial economy. The development gap between the east and the central and western regions has become more and more obvious, and the development factors are constantly changing from the central and western regions. The region shifted to the eastern region [4, 5].

People are the main body of economic activities, and labor is the main economic factor [6, 7]. With the rise of the eastern region, people will inevitably choose to move to economically developed regions in order to seek better jobs and higher incomes. Because of this, with the relaxation of the population migration policy, China's domestic labor-based population migration has become more frequent. Large-scale young labor-based populations are constantly migrating from rural areas to cities, and from central and western regions to eastern regions. It has become the unprecedented and largest population migration activity in human history. It has also surpassed the population factors dominated by birth and death, and has become an important factor in the development of Chinese cities and social economy [8, 9]. Therefore, the relationship between population changes and economic growth has always been the focus of scholars, and it is also one of the most basic starting points for studying all population issues. Large-scale population migration is one of the most prominent facts in China's current economic and social development, and has a profound impact on both population changes and economic development. Carry out in-depth research on the internal mechanism between population migration, population and economic production activities in China, and provide practical experience for the development of cities in China and the world. At the same time, the coordination between population and urban economy is also the basis for realizing regional sustainable development, and it is a fundamental issue that constitutes a country's economic stability.

Grasping the relationship between population mobility and economic development between regions is helpful to the design and improvement of urban planning and construction, infrastructure investment and basic public services, and points out the direction for China's current urbanization development.

With the acceleration of China's urbanization process, urban agglomeration gradually replaces single city as the power source to promote regional development [10]. Both world-class urban agglomerations and the emerging urban agglomerations in China have an important impact on the economic growth and association development of a country and play a central role in regional development pattern and international competition. Since 2006, when China first proposed urban agglomeration as the main body to promote urbanization and gave play to the leading role of central cities, China has gradually formed the embryonic form of urban agglomeration represented by the Yangtze River Delta, Pearl River Delta and Beijing-Tianjin-Hebei urban agglomeration. The proportion of China's urban agglomeration GDP in the country has increased from 70.56% in 2006 to 88.10% in 2017, while the proportion of the total population of urban agglomerations in the total population of the country has risen from 61.21% to 75.30%. Among them, the Yangtze River Delta urban agglomeration, one of the world-class urban agglomerations, is not only an important growth pole for promoting the economic development of the Yangtze River Delta region and the whole country, but also the region with the highest urban density and the most densely populated population in our country. From 2000 to 2010, the floating population of the Yangtze River Delta urban agglomeration increased from 21.69 million to 47.83 million. Although the scale of the migrant population has slowed down in recent years, the total migrant population is still at a relatively high level. The floating population can change the structure of the labor force in the region, optimize the allocation of resources, and reshape the pattern of economic growth [11].

By combing the literature related to urban agglomerations, population mobility and economic growth, we use panel data models to conduct empirical analysis on population mobility to promote urban development. This study supplements the existing literature through the following three aspects: (1) Taking the Yangtze River Delta urban agglomeration, which has the most obvious urban development and population mobility, as the research object, and integrating the panel data of all prefecture level cities in the urban agglomeration to carry out empirical analysis, in order to explore the impact of population mobility on urban economic development, and to provide a supplement for the study of regional economy and urban population. (2) Through the influence path of data population mobility on economic growth, this part puts forward the hypothesis of transmission mechanism, and uses a series of empirical analysis to verify the hypothesis. At the same time, the results pass the robustness and endogenous test. (3) Considering the difference of urban development within the triangle city group, we have discussed the influence of population mobility of different scale cities on urban development.

Therefore, this chapter takes the Yangtze River Delta (YRD) urban agglomeration as the research object, adopts the panel data model, and empirically analyzes the mechanism of regional population mobility to promote urban development at the meso level.

This chapter is organized as follows. In Chapter 5.2, the variables, methods and the materials are introduced. The general information of urban development and population mobility in YRD and empirical analysis of population mobility promoting urban development is presented in Chapter 5.3. Chapter 5.4 is the transmission mechanism and robustness test of population mobility promoting urban development. The Chapter 5.5 is the summary.

5.2 Study area, methods and materials

5.2.1 Study areas

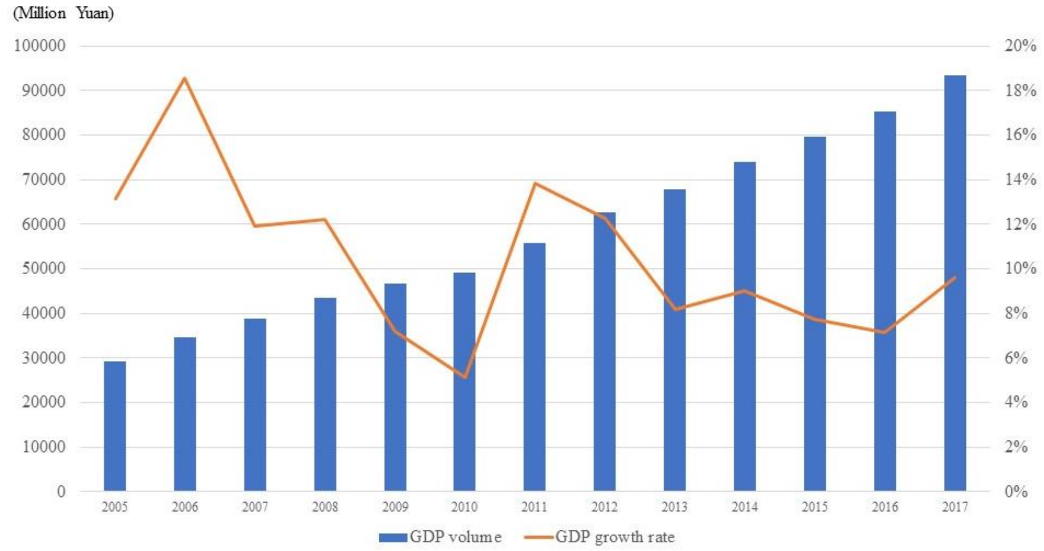


Fig. 5-1 GDP and growth rate of YRD Urban Agglomeration from 2005 to 2017.

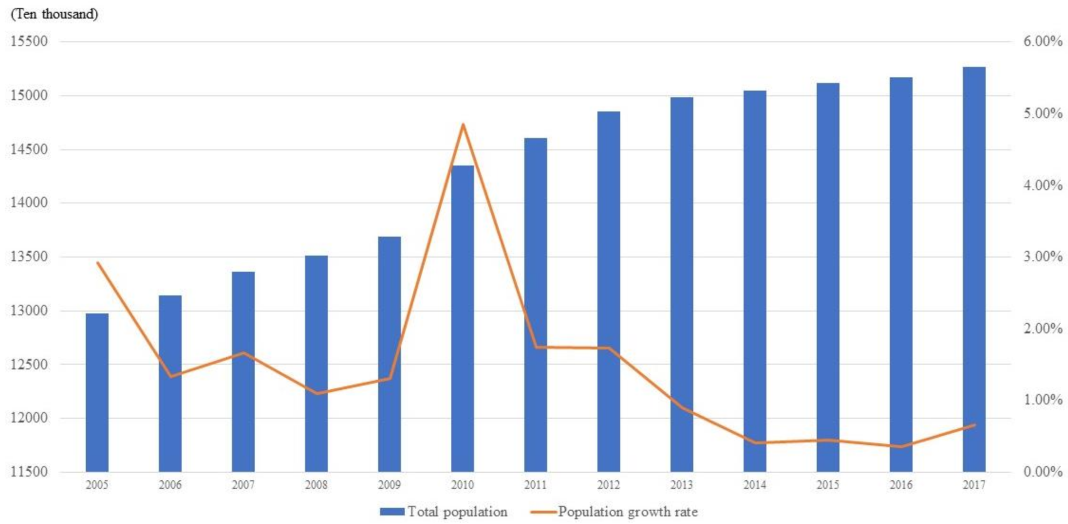


Fig. 5-2 Population volume and growth rate of YRD Urban Agglomeration from 2005 to 2017.

In this chapter, taking the Yangtze River Delta (YRD) urban agglomeration as the research object, the mechanism of regional population mobility to promote urban development is studied at the regional scale. The YRD urban agglomeration is an important intersection between the "Belt and Road" and the Yangtze River Economic Belt. It has a pivotal strategic position in the overall situation of China's national modernization and opening up. The YRD urban agglomeration is the most developed urbanized region with the highest degree of urban agglomeration in China. The YRD is one of the most dynamic regions in China's economic development. It accounts for only 2.1% of China's regional area, accounts for one quarter of China's total economic output and more than one quarter of China's industrial added value, and is regarded as an important engine of China's economic development. It is the most developed region in China. Fig. 5-1 and Fig. 5-2 show the GDP and total population of the Yangtze River Delta Urban Agglomeration in 2005-2017, respectively.

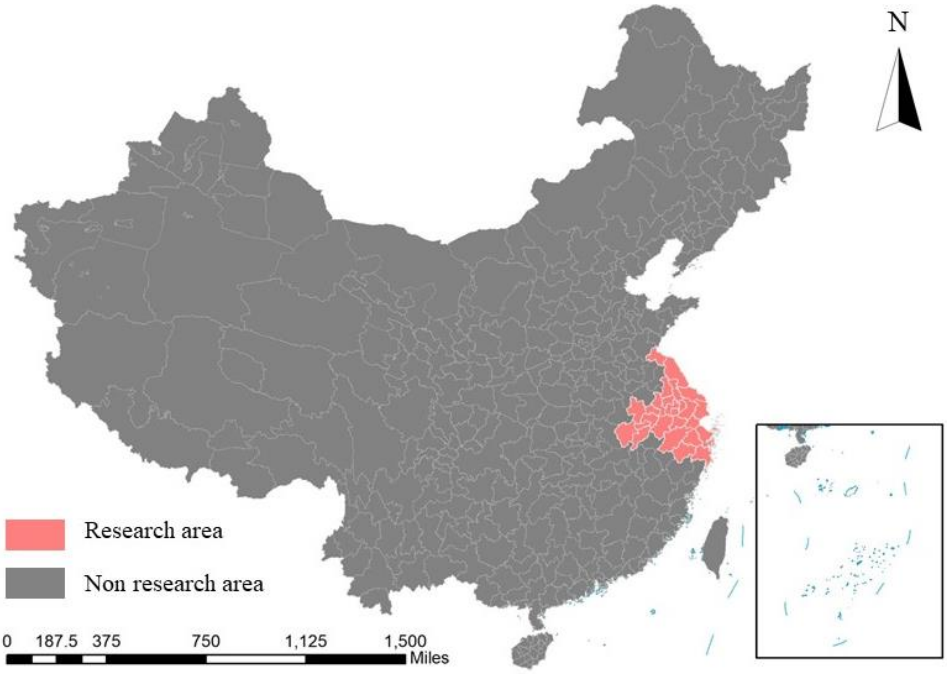


Fig. 5-3 Location of the research area.

The YRD urban agglomeration, with Shanghai as the center, located in the east China region, including: Shanghai, Nanjing, Wuxi, Changzhou, Suzhou, Nantong, Yancheng, Yangzhou, Zhenjiang, Yangzhou, Taizhou, Hangzhou, Ningbo, Jiaxing, Huzhou, Shaoxing, Jinhua, Zhoushan, Taizhou, Anhui Hefei, Wuhu, Maanshan, Tongling and Anqing, Chuzhou, Chizhou, Xuancheng, a total of 26 cities, with area of 21.17 square kilometers, about 2.2% of Chinese area. In addition, according to the city size classification standard formulated by the State Council in 2014, we divided 26 cities into large cities, small and medium sized cities. Table 5-1 is the city classification.

Table 5-1. City classification in the YRD urban agglomeration.

City classification	cities
Mega cities (Urban resident population > 10 million)	Shanghai
Large cities (Urban resident population > 5 million)	Nanjing, Hangzhou, Suzhou,
Medium sized cities (Urban resident population > 1 million)	Wuxi, Changzhou, Wuhu, Hefei, Taizhou (Zhejiang), Shaoxing, Jiaxing, Taizhou (Jiangsu), Zhenjiang, Yangzhou, Yancheng, Nantong, Jinhua, Ningbo
Small cities Urban resident population < 1 million	Anqing, Xuancheng, Chizhou, Chuzhou, Tongling, Maanshan, Zhoushan

5.2.2 variable selection and research data

Since the reform and opening up, China's large-scale population mobility has promoted the process of urbanization and industrialization, and affected the mode and pattern of economic development [12, 13]. Population mobility is the result of a combination of many factors. At the same time, foreign scholars Lewis's dual economic structure model, G.W. Jogenson model and M.P. Todro model all analyze the flow of labor between industries and believe that the flow of labor between urban and rural areas is an important way to upgrade the industrial structure [14, 15]. Moving from the agricultural sector to the secondary and tertiary industries can promote the upgrading of the industrial structure [16, 17].

From the perspective of empirical research, domestic scholars have also studied the relationship between population flow and industrial structure upgrading, and found that the cross regional flow of population promotes the flow of labor force from the primary industry to other industries, promotes the matching of labor supply and demand, changes the employment structure and industrial structure, further makes the industry obtain structural dividend, thus promoting the optimization and upgrading of industrial structure [18, 19]. At the same time, a large number of literature studies show that the optimization and upgrading of industrial structure has a positive impact on economic growth [20-22]. The YRD region, as a highly concentrated area of population and industry, has a large number of foreign population and foreign investment, which can effectively realize the combination of labor and capital, promote the rapid development of service industry, promote the industrial structure to present a high-level trend, and promote the smooth operation of the economy. Therefore, we can put forward such a hypothesis: population mobility can promote the upgrading of industrial structure of YRD urban agglomeration, thus promoting economic growth.

According to the theory of endogenous economic growth, human capital promotes economic growth through technological innovation and knowledge spillover, and population mobility as well as aggregation, as the key factors affecting human capital, have an important impact on urban scientific and technological innovation [23, 24]. In recent years, the human capital of the floating population in China has been continuously improved, and the average number of years of education of the floating population has increased significantly [25]. The flow of population provides a large number of scientific and technological talents and scientific research and innovation talents for the inflow areas, which can significantly improve the local labor structure, increase the stock of human capital, and promote the agglomeration of knowledge and technology, thus promoting the independent innovation of enterprises and enhancing the innovation ability of cities [26, 27]. Scientific and technological innovation is an important driving force for the economic development of a country and a region, and plays an obvious role in improving the quality and efficiency of economic growth. The Yangtze River Delta region has a high level of innovation cooperation and industrial chain division, rich scientific and technological innovation resources [28]. Therefore, we put forward the second hypothesis: population mobility can promote scientific and technological innovation and knowledge spillover in the YRD region, thus promoting economic growth [29].

Whether absolute income theory, relative income theory or permanent income theory, all think that income is the basis of consumption and an important factor to determine the level of residents' consumption. The floating population has the dual characteristics of production and consumption, especially in recent years, the trend of floating population's family has an important impact on the consumption level and structure of residents [30]. The migration and aggregation of population can not only stimulate the consumption of rural residents, but also accelerate the division of labor and the improvement of productivity, so as to increase the income of residents and stimulate the consumption level [31, 32]. In addition, consumption, as one of the "troikas" driving economic growth, not only plays an important role in economic development, but also promotes the floating population to integrate into the local society [33, 34]. Due to the developed economic level and high-quality material living standard, the YRD urban agglomeration has attracted a large number of floating populations to enter the local area through "voting with feet" [35]. Therefore, we put forward a third hypothesis: population mobility can promote residents' consumption, thus promoting economic growth.

Floating population is not only the source of urban labor supply, but also the competitor of urban employment. The relationship between the floating population and the local population affects the attitude and decision-making ideas of urban managers towards the floating population, and then affects the sustainable and stable development of the urban labor market in the future [36, 37]. The common view is that the floating population and the local population are competitive. The floating population has occupied the employment opportunities of the local population, which leads to the employment difficulties of the local population and the increase of the unemployment rate, especially the labor groups with low education and low skills. However, from the economic point of view, the floating population expands the scale of the labor market, promotes the redivision of labor in the labor market, and improves the efficiency of the labor market, which belongs to the "Pareto improvement" effect [38-40]. Therefore, we propose a fourth hypothesis: population mobility can improve the regional labor market and employment, thus promoting economic growth.

Therefore, the indicators we select in this chapter include gross regional product, population mobility scale, urban industrial structure, urban science and technology innovation level, urban residents' consumption level and labor market employment scale [41]. In addition, considering other factors affecting regional economic development, we also control a series of urban development indicators, including the level of opening up, foreign investment, urbanization rate, investment in fixed assets, disposable income of residents and total investment in urban infrastructure. Among them, the dependent variable is GRP which can represent the level of urban economic development. We use the difference between the total population at the end of the year and the number of registered residence population to represent the scale of city's population mobility, which is positive for the net population inflow, and the negative value is the net outflow of population. At the end of the year, the added value of the tertiary industry is used to represent the change of urban industrial structure. Urban innovation index is used to comprehensively measure the level of scientific and technological innovation level of a city. The per capita consumption expenditure of urban residents and the new jobs in the labor market at the end of the year are used as the indicators to measure the consumption level and employment market respectively. All data come from China City Statistical Yearbook, China City Construction Yearbook, National Economic and Social Statistics Bulletin. Meanwhile, we acquire the supplemental data from provincial and prefectural Statistical Yearbooks. Table 5-2 is the description of all variables in this chapter.

Table 5-2. Description of all variables in this chapter.

Class	Variable	Notation	Explanation
Dependent variable	Gross regional product	GRP	Gross regional product at the end of the year
Independent variable	Population mobility scale	MIG	The difference between the total population and registered residence population at the end of the year
	Industrial structure	IND	The added value of the tertiary industry at the end of the year
	Science and technology innovation level	TEC	Urban innovation index comprehensively quantifies the level of scientific and technological innovation level
	residents' consumption level	COS	Total annual consumption expenditure of urban residents at the end of the year
	labor market employment scale	EMP	New jobs in the labor market at the end of the year
Control variable	Opening up level	OPE	Total import and export volume at the end of the year
	Foreign capital participation	FCP	Total foreign investment at the end of the year
	Urbanization level	URB	The proportion of urban population in total population at the end of the year
	Government size	GOV	Total investment in fixed assets at the end of the year
	Resident income level	RIL	Disposable income of urban residents at the end of the year
	Urban construction scale	UCS	Infrastructure investment at the end of the year

5.3 Empirical analysis of population mobility promoting urban development

Fig. 5-4 to 5-7 show the population mobility scale and economic growth of all cities in the YRD urban agglomeration from 2005 to 2017. Among the mega cities, Shanghai has the largest economic aggregate. Although the scale of Shanghai's inflow population has been stable at 9-10 million since 2010, the total economic volume has shown a stable growth trend and the growth rate is obvious. Among the large cities, the net inflow of population in Hangzhou, Nanjing and Suzhou is also stable. Among them, the total net inflow of population in Suzhou is the largest, and it continues to be stable at about 4 million. Meanwhile, the economic volume of Suzhou is the largest of the three large cities.

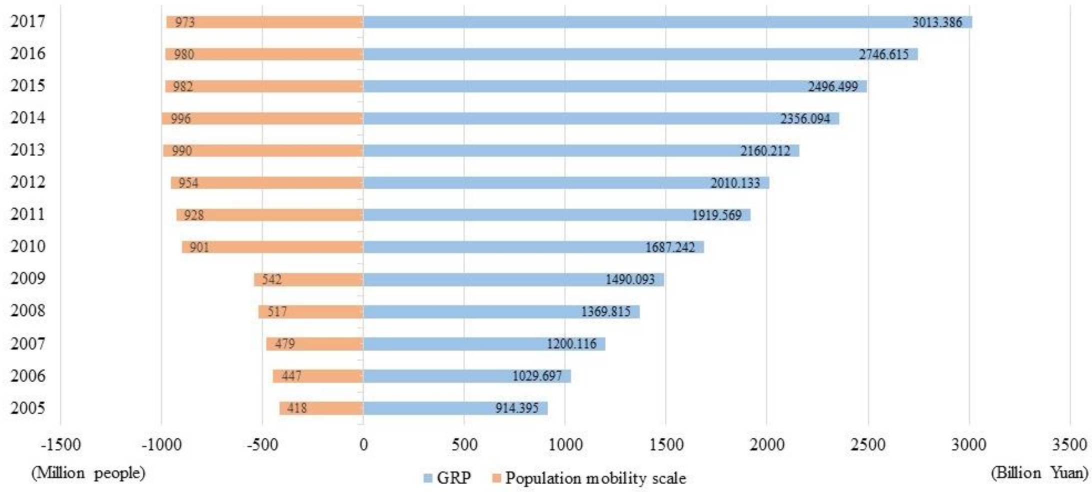


Fig. 5-4 The scale of population mobility and economic development with mega cities (Shanghai) in YRD urban agglomeration from 2005 to 2017 (A positive value indicates a net population inflow, a negative value indicates a net population outflow).

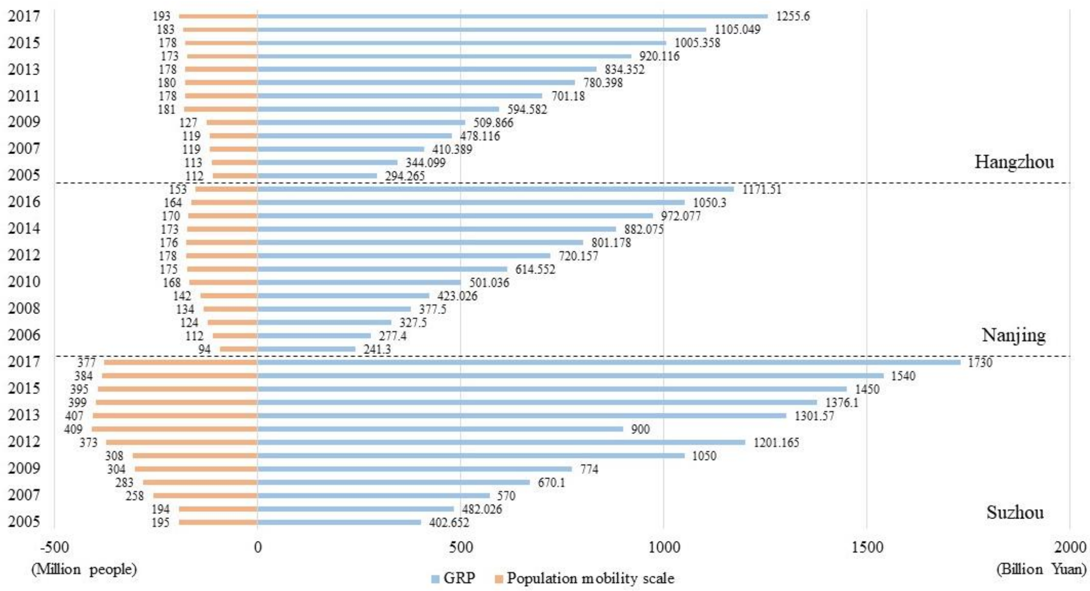


Fig. 5-5 The scale of population mobility and economic development with large cities in YRD urban agglomeration from 2005 to 2017 (A positive value indicates a net population inflow, a negative value indicates a net population outflow).

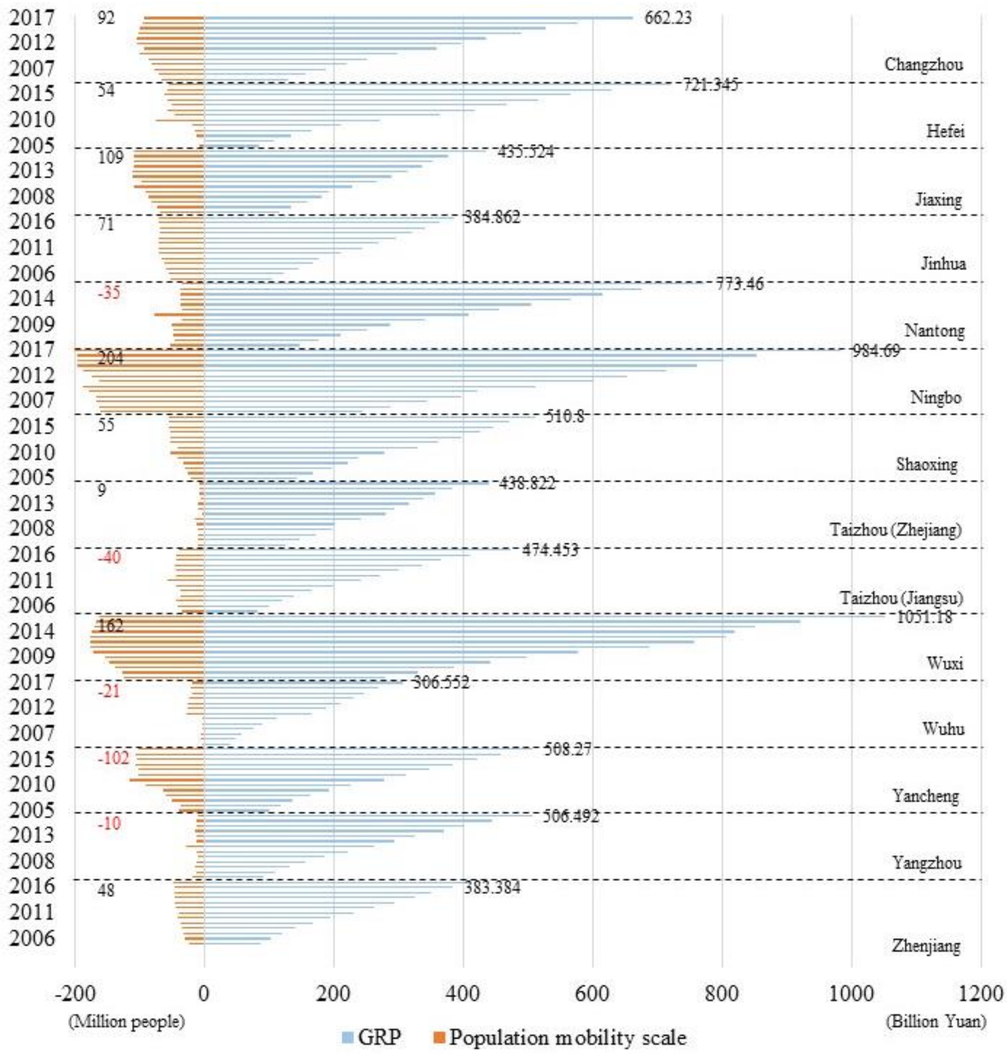


Fig. 5-6 The scale of population mobility and economic development with medium-sized cities in YRD urban agglomeration from 2005 to 2017 (A positive value indicates a net population inflow, a negative value indicates a net population outflow).

The net population inflow scale of Nanjing and Hangzhou is about 1.5 million, and the economy is also showing an upward trend year by year. However, in medium-sized cities, there are obvious changes in population mobility. The scale of floating population in Nantong, Taizhou, Wuhu, Yancheng and Yangzhou is negative, which indicates that the above cities continue to face the problem of net population outflow from 2005 to 2017. Among them, Yancheng has a net population outflow scale of more than 1 million, becoming the largest city among medium-sized cities. Although the economy of the above-mentioned cities shows a rising trend year by year, there is still a gap between the total economic volume and the average annual growth rate compared with other cities. Among all the small cities, only Zhoushan has a positive population flow scale, which indicates that all cities except Zhoushan are facing the problem of net population loss from 2005 to 2017. Anqing has a net population loss scale of more than 500000, becoming the most serious one among small cities. Although the economy of the above cities is growing year by year, the overall economic volume and average annual growth rate are small.

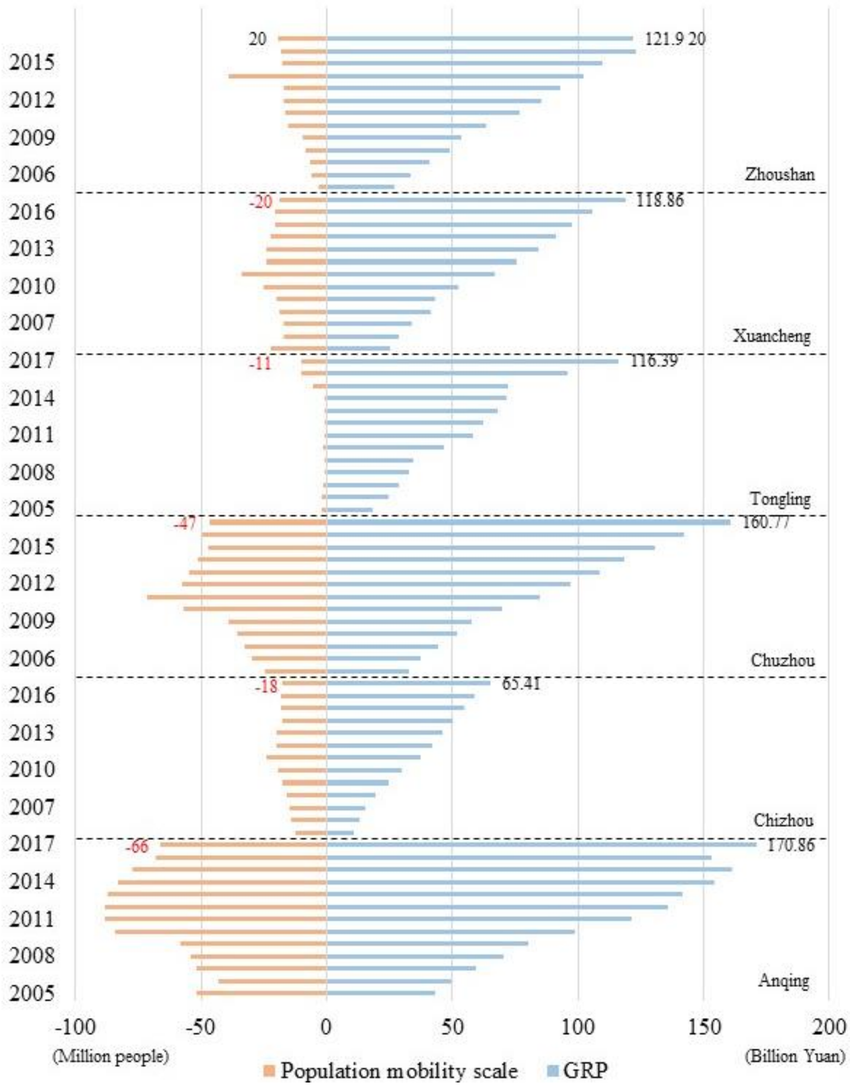


Fig. 5-7 The scale of population mobility and economic development with medium-sized cities in YRD urban agglomeration from 2005 to 2017 (A positive value indicates a net population inflow, a negative value indicates a net population outflow).

In order to investigate the impact of population mobility in the YRD urban agglomeration on economic growth, we take the scale of population mobility as a core dependent variable and use panel data regression models for empirical analysis. For the detailed regression model in this chapter, please refer to the methodology in Chapter 2. In addition, by analyzing the transmission mechanism of population mobility to regional economic growth, we put the interaction items of population flow and industrial structure, population flow and scientific and technological innovation, population flow and resident consumption, population flow and market employment into the benchmark model, and get the extended model of empirical analysis.

Since the panel regression model includes pooled model, random effect (RE) model and fixed effect (FE) model, we carry out pooled regression, random effect regression and fixed effect regression on the basis of the benchmark model, and the results are shown in Table 6-3. In the three models with different effects, the independent variables and some control variables are significant.

Therefore, in order to determine the regression model, we conduct F-test and Hausman test on the benchmark model. The benchmark model has passed F-test and Housman test at 1% significance level, that is, refusing to accept the pooled model and random effect model. Therefore, the fixed effect model in the panel data model is adopted in this chapter for empirical analysis.

In the analysis of the fixed effect regression model, the coefficient of population mobility scale is positive, and has passed the 1% significance level test, which shows that the population flow has a significant positive role in promoting the economic growth in Yangtze River Delta region. Every 1% increase in population mobility will increase the GRP by 0.259%. In addition, as far as the control variables are concerned, the coefficient of opening up level is positive, indicating that the increase in the total import and export volume of the Yangtze River Delta region will promote economic growth. The prosperity of import and export trade can not only promote the redistribution of resources, but also increase labor productivity, and at the same time alleviate employment pressure and provide more jobs. According to the internationally accepted calculation method, every 100 million U.S. dollars' worth of products exported can provide about 10,000 jobs. The coefficient of foreign capital participation is positive and passed the test at the 5% significance level, indicating that foreign capital participation has a positive effect on the economic growth of the Yangtze River Delta region. As an important source of capital, foreign capital participation can not only improve the quality of investment, but also promote the technological progress and industrial structure upgrade of enterprises, thereby achieving economic growth. The regression coefficient of urbanization level is not significant, which indicates that the improvement of urbanization level in Yangtze River Delta region will not bring economic growth. The possible reason is that in the process of development, the excessive pursuit of development speed and the neglect of development quality lead to the mismatch between urbanization level and regional economic development. The coefficient of residents' income level is positive, and has passed the significance test of 1% level, which shows that the improvement of residents' income level in Yangtze River Delta region promotes economic growth. The increase of residents' income has led to the improvement of consumption level, increased the consumption of domestic products, expanded domestic demand, thus creating more jobs, improving the labor market, especially the developed service industry in the Yangtze River Delta region, thus promoting the regional economy. The coefficient of urban construction scale is negative and significant, which indicates that the construction scale of urban infrastructure has inhibitory effect on economic growth. The possible reason is that the construction of transportation infrastructure occupies the main position in the scale of urban construction, and the convenience of transportation accelerates the population mobility among regions, thus affecting the growth of urban economy. The coefficient of government size is positive and significant, which shows that the increase of fixed assets investment improves the regional economy. On the one hand, fixed assets investment improves the future production and service capacity, thus affecting the labor employment market. On the other hand, it promotes regional consumption and stimulates economic growth.

Table 5-3. Benchmark regression and panel model selection.

Variables	Benchmark model		
	Pooled model	FE model	RE model
Migration scale	0.198	0.259	0.252
	(6.885**)	(7.561**)	(8.280**)
Control variables			
Opening up level	0.258	0.219	0.233
	(9.578**)	(7.876**)	(8.768**)
Foreign capital participation	0.187	0.093	0.134
	(5.738**)	(2.562*)	(4.010**)
Urbanization level	-0.030	-0.008	-0.009
	(-1.901)	(-0.409)	(-0.532)
Government size	0.456	0.418	0.442
	(16.271**)	(12.019**)	(14.117**)
Resident income level	0.099	0.155	0.121
	(5.553**)	(6.141**)	(5.770**)
Urban construction scale	-0.060	-0.087	-0.076
	(-2.305*)	(-3.351**)	(-3.065**)
Constant	0.068	0.105	0.006
	(2.211*)	(2.776*)	(0.105)
R^2	0.958	0.846	0.910
F-test	F (25,305) =6.165, p=0.000		
BP-test	$\chi^2(1)$=113.155, p=0.000		
Hausman test	$\chi^2(7)$=14.917, p=0.037		

Note. t value in parentheses.

* p<0.05.

** p<0.01.

5.4 Transmission mechanism and robustness test of population mobility promoting urban development

5.4.1 Transmission mechanism of population mobility promoting urban development

In order to further study the transmission mechanism of population mobility on economic growth in the Yangtze River Delta region, we put the interaction of population mobility and industrial structure, population mobility and household consumption, population mobility and scientific and technological innovation, population mobility and labor market into the extended model, and the empirical results of the extended model are obtained. Table 6-4 shows the summary results of the extended model after adding interaction items.

Table 5-4. An empirical analysis on the transmission mechanism of regional economic growth affected by population mobility.

Variables	Extended model			
	Industrial structure	Resident consumption	Science and technology innovation	Employment market
Migration scale	0.026 (2.748*)	0.519 (15.364**)	0.108 (4.469**)	0.065 (2.887*)
Migration scale×Industrial structure	0.512 (11.472**)			
Migration scale×Resident consumption		0.898 (27.452**)		
Migration scale×Science and technology innovation			0.335 (19.373**)	
Migration scale×Employment market				0.264 (2.966**)
Control variables				
Total export-import volume	0.038 (1.371)	-0.064 (-3.544**)	-0.019 (-0.838)	0.191 (6.622**)
Foreign capital	0.023 (0.744)	-0.110 (-5.248**)	-0.081 (-3.098**)	0.061 (1.634)
Urbanization rate	0.009 (0.571)	0.012 (1.101)	-0.004 (-0.275)	-0.012 (-0.617)
Fixed assets investment	0.467 (15.894**)	0.444 (23.758**)	0.457 (19.535**)	0.430 (12.435**)
Disposable income	0.144 (6.831**)	0.048 (3.430**)	0.088 (5.082**)	0.153 (6.118**)
Infrastructure investment	-0.191 (-8.114**)	-0.034 (-2.424*)	-0.030 (-1.694)	-0.079 (-3.059**)
Constant	0.086 (12.111**)	0.032 (2.219*)	0.139 (8.727**)	0.106 (5.508**)
R^2	0.914	0.961	0.939	0.867

Note. t value in parentheses.

* $p<0.05$.

** $p<0.01$.

From the regression results of the extended model, we can find that population mobility affects regional economic growth by optimizing industrial structure, increasing residents' consumption, promoting technological innovation and changing labor employment. Specifically, in the optimization model of industrial structure, the regression coefficient of population mobility is positive, and the coefficient of interaction between population mobility and industrial structure is also positive, and has passed the significance test at 1% level, which shows that population mobility between regions can optimize the allocation of labor factors, promote the transformation and upgrading of industrial structure, thus promoting the growth of regional economy. In the model of science and technology innovation, population mobility and interaction are positive and significant, and pass the significance test of 1% level, which shows that population mobility provides talent guarantee for scientific and technological innovation, improves the level of scientific and technological innovation of cities, and promotes the high-quality development of regional economy. In the model of residents' consumption, the coefficients of population mobility and interaction are both positive and significant, which indicates that population inflow improves residents' consumption level and promotes consumption upgrading, thus positively affecting the growth of urban economy. In the model of employment market, the coefficients of population mobility and interaction are both positive and significant, which indicates that population mobility does not bring pressure to the local labor market. On the contrary, population mobility expands the scale of local labor market, promotes the redivision of labor and improves the efficiency of labor market. Therefore, population mobility can also positively affect economic growth by promoting the prosperity of the labor market.

In this part, we empirically analyze the transmission mechanism of population mobility promoting economic development. The regression results show that population mobility can promote regional economic growth by optimizing industrial structure, increasing residents' consumption, strengthening scientific and technological innovation and adjusting employment market.

5.4.2 Robustness test of population mobility promoting urban development

After adding the interaction item to the benchmark model, the direction of the core independent variable—population mobility is consistent with the benchmark model, indicating that the results of the benchmark model in this chapter are robust. On this basis, we further use other methods to test robustness. Considering that the impact of population mobility on economic growth may have a certain time lag, we put the population mobility lag phase into the benchmark and extended model instead of the population mobility index, and conduct regression respectively. Since the unpredictable factors affecting the current economic growth will not affect the last population mobility, this method can effectively reduce the coherence between random interference items and core independent variables, and to some extent avoid the interference of endogenous on regression results. See table 5-5 for detailed results of robustness test.

Table 5-5. Robustness test of benchmark model and extended model.

Variables	Lagged model			
	Industrial structure	Resident consumption	Science and technology innovation	Employment market
Lagged migration scale	0.018 (2.181*)	0.416 (11.216**)	0.164 (6.628**)	0.087 (2.212*)
Lagged migration scale× Industrial structure	0.349 (9.282**)			
Lagged migration scale× Resident consumption		0.699 (18.893**)		
Lagged migration scale× Science and technology innovation			0.428 (11.206**)	
Lagged migration scale× Employment market				0.336 (3.316**)
Control variables				
Total export-import volume	0.022 (1.715)	-0.072 (-2.332*)	-0.032 (-1.257)	0.189 (8.829**)
Foreign capital	0.102 (0.661)	-0.091 (-6.672**)	-0.066 (-2.231*)	0.669 (1.128)
Urbanization rate	0.006 (1.212)	0.034 (1.258)	-0.012 (-0.288)	-0.207 (-0.207)
Fixed assets investment	0.121 (9.646**)	0.639 (12.127**)	0.551 (16.625**)	0.662 (10.219**)
Disposable income	0.215 (7.756**)	0.105 (8.827**)	0.096 (7.721**)	0.225 (4.428**)
Infrastructure investment	-0.302 (-10.108**)	-0.012 (-2.667*)	-0.130 (-1.891)	-0.901 (-2.224*)
Constant	0.071 (8.816**)	0.064 (2.501*)	0.248 (7.726**)	0.307 (6.652**)
R^2	0.826	0.905	0.899	0.923

Note. t value in parentheses.

* $p<0.05$.

** $p<0.01$.

The results of robustness test show that for the benchmark model and the extended model with interactive items, the direction and significance of the lag term and the interactive term of the core explanatory variables are consistent with the previous analysis, and other control variable coefficient and significance also did not change significantly, the lagged population mobility scale still has a promoting effect on economic growth, which shows that the empirical results of this chapter are robust.

5.4.3 Heterogeneity analysis of population mobility promoting urban development

In order to further investigate the heterogeneous impact of population mobility on economic growth, we divide all cities into two categories, namely, large cities, small and medium-sized cities. Table 5-6 shows the results of heterogeneity analysis.

Table 5-6. Analysis of the Heterogeneity of population mobility on economic growth.

Variables	Heterogeneous model	
	Large cities	Small and medium-sized cities
Migration scale	0.047	0.071
	(2.847**)	(1.181)
Control variables		
Opening up level	0.243	0.770
	(6.491**)	(13.027**)
Foreign capital participation	0.172	-0.055
	(6.611**)	(-1.193)
Urbanization level	0.064	0.007
	(1.571)	(1.553)
Government size	0.333	0.230
	(7.959**)	(14.096**)
Resident income level	0.285	0.016
	(8.471**)	(2.300*)
Urban construction scale	-0.096	-0.034
	(-3.097**)	(-3.301**)
Constant	0.047	-0.239
	(2.847**)	(-7.916**)
R^2	0.891	0.953

Note. t value in parentheses.

* $p<0.05$.

** $p<0.01$.

The regression results show that population mobility has a positive effect on urban economic growth, but this effect is only significant in large cities, but not in small and medium-sized cities. This shows that large cities have a strong economic agglomeration effect, and it is easier to attract the inflow of population and combine with their own advantageous resources to promote economic development. Small and medium-sized cities have a single economic structure and a small scale of economic aggregate. The lack of diversified economic industries limits the attraction of population inflow, resulting in the impact of population inflow on economic growth is not obvious.

5.4 Summary

In this chapter, we have collected a series of urban development indicators from 2005 to 2017, including regional GRP, industrial structure changes, residents' consumption levels, technological innovation capabilities, labor market employment, total imports and exports, foreign capital participation, fixed asset investment, etc., taking the Yangtze River Delta urban agglomeration as the research object, and using the panel data regression model to explore the role of regional population mobility on the economy growth.

Our findings are as follows: (1) The population mobility between regions has an important impact on economic growth. The net influx of urban population can greatly promote economic development. As far as the Yangtze River Delta is concerned, for every 1% increase in the size of the city's net population inflow, their total GRP will increase by 0.259%. (2) The net influx of population between regions can change the industrial structure of the city, promote the development of the city to the tertiary industry, and thus affect economic growth. The net influx of urban population can also affect household consumption, expand domestic demand, and stimulate economic growth. The input of labor has changed the age structure of the city. High-tech talents and professionals have enhanced the city's scientific research and innovation capabilities, laid a good foundation for the development of high-tech industries, and affected economic development in the long run. Migrant labor has changed the local job market, expanded the scale of the job market, and promoted the division of labor and specialization, making the overall labor participation rate and employment efficiency of the city higher and ensuring economic development. (3) The effect of population mobility on economic growth is heterogeneous. Compared with small and medium-sized cities, population mobility in large cities has a more significant impact on economic growth. Generally speaking, due to their own development constraints, it is difficult for small and medium cities to attract a large influx of population. This is consistent with our analysis. At the same time, small and medium cities are difficult to form scale effects and have limited capacity to absorb population. If things go on like this, the development of the city will fall into an endless cycle. The shortage of labor restricts the economic development, and the slow economic development makes the city lack sufficient attraction to the floating population, and eventually develop into a shrinkage city.

Meanwhile, there are still some limitations in this chapter. First, we use the economic growth of the city as the proxy of urban development, there is a certain deviation. Urban development is all-round, and economic growth is only one aspect of urban development. The development of cities at the social level, such as the environment, public services, and social resource supply, are not mentioned in this chapter. However, we believe that the economic prosperity of a city is the foundation of all development. Only when the economy advances steadily can there be sufficient financial resources to support the supply of social resources in the city. In the following research, we will try to explore the connection between population mobility and urban social development. Second, there is insufficient analysis of the heterogeneous impact of population mobility on promoting economic growth. In recent years, the strict household registration management system has restricted the free mobility of population. The key tasks of the new urbanization construction issued by the State Council in 2019 emphasize the full liberalization and relaxation of population free settlement, which provides new impetus for urban development. Therefore, the analysis of the heterogeneity of population mobility in promoting urban development should involve more aspects, such as areas with strict household registration management and loose household registration management, areas with higher education and areas with lower education. Third, we only selected areas with obvious population mobility and the most developed economy as the research objects, which will inevitably cause errors in the results of the empirical analysis. In the future research, we try to take all prefecture-level cities as the research objects, and systematically analyze the two-way connection between population mobility and urban development.

Reference

1. Cao, Y., Research on the Contribution of Beijing's Permanent Residents to Economic Growth Based on C-D Production Function. *Business Economic Research* **2015,** 19.
2. Li, Y.; Yin, J., High savings, high investment and China's economic growth in the process of labor transfer. *Economic Research* **2005,** 02.
3. Wang, Z., Floating Population and Economic Development——Research Based on the Data of Prefecture Level City. *Modern Urban Research* **2013,** 03.
4. Yang, S.; Gao, X., Research on the Impact of Migrant Workers on the Economic Development of the Places of Inflow——Taking Shanghai as an Example. *Economic Reform* **2012,** 06.
5. Zhang, L., Analysis of the Contribution of Floating Population to Urban Economy: Taking Shanghai as an Example. *Population Research* **2015,** 04.
6. Ma, X.; Zhang, Y., A Preliminary Study on the Quantification of the Economic Contribution of Urban Floating Population. *Population Research* **2004,** 04.
7. Yuan, X., Research on the Contribution of Floating Population to China's Economic Growth. *China Prices* **2017,** 03.
8. Du, Y.; Cai, F.; Qu, X.; Cheng, J., Continuing the Chinese miracle: reaping dividends from the reform of the household registration system. *Economic Research* **2014,** 08.
9. Li, Z., Research on the Contribution of Floating Population to the Economic Growth and Fiscal Revenue of Big Cities——Based on the Calculation of Ningbo City. *Shandong Association Forum* **2017,** 06.
10. Fang, C., Significant progress and future development directions in the study of Chinese urban agglomerations. *Acta Geographica Sinica* **2014,** 69.
11. Yin, J.; Li, H., Retrospect and Prospect of China's Population mobility and Urbanization Process. *Urban Problem* **2012,** 12.
12. Wang, F.; Wei, X.; Liu, J.; He, L.; Gao, M., Impact of high-speed rail on population mobility and urbanisation: A case study on Yangtze River Delta urban agglomeration, China. *Transportation Research Part A: Policy and Practice* **2019,** 127, 99-114.
13. Wu, J.; Yu, Z.; Wei, Y. D.; Yang, L., Changing distribution of migrant population and its influencing factors in urban China: Economic transition, public policy, and amenities. *Habitat International* **2019,** 94, 102063.
14. Jorgenson, D. W., Surplus agricultural labour and the development of a dual economy. *Oxford economic papers* **1967,** 19, (3), 288-312.
15. Lewis, W. A., Economic development with unlimited supplies of labour. **1954**.
16. Kuznets, S., Fertility differentials between less developed and developed regions: Components and implications. *Proceedings of the American Philosophical Society* **1975,** 119, (5), 363-396.
17. Pennock, A., The political economy of domestic labor mobility: Specific factors, landowners, and education. *Economics & Politics* **2014,** 26, (1), 38-55.
18. Deng, Z.; Dan, B., On the transfer of rural surplus labor and the evolution of regional industrial structure. *China's Rural Economy* **2005,** 8.
19. Zhao, N., Research on the spatial effect of labor flow and industrial structure adjustment. *Statistical Research* **2016,** 33.
20. Gan, C.; Zheng, R., Research on the evolution of industrial structure and productivity growth since reform and opening-up: A test of China's "structural dividend hypothesis" from 1978 to

2007. *Chinese Industrial Economy* **2009,** 2.

21. Jing, X., Advanced industrial structure and economic growth——An empirical analysis of the Yangtze River Delta region. *Journal of Nantong University: Social Science Edition* **2005,** 3.
22. Yan, H., Advanced industrial structure and its influencing factors in the Yangtze River Delta region. *Financial Science* **2010,** 12.
23. Nelson, R. R.; Phelps, E. S., Investment in humans, technological diffusion, and economic growth. *The American economic review* **1966,** 56, (1/2), 69-75.
24. Xu, N.; Lv, L.; Huang, R.; Hu, H., China's urban population migration and innovation. *Regional Research and Development* **2016,** 35.
25. Duan, C.; Xie, D.; Lv, L., Migration of China's population. *Population Research* **2019,** 43.
26. Huang, R.; Liang, Y.; Lv, L., The relationship between urban population structure and innovation ability—— An empirical analysis based on Chinese cities. *urban Studies* **2014,** 21(9), 84-91.
27. Wang, Y.; Zhang, G., Population agglomeration, communication externalities and independent innovation of enterprises. *Financial and Trade Economy* **2015,** 5, 132-146.
28. Hong, Y., On innovation driven economic development strategy. *Economist* **2013,** 1, 5-11.
29. Yin, D.; Wu, H.; Jin, Z., Innovation network, knowledge spillover and high quality integrated development: evidence from Yangtze River Delta urban agglomeration. *Shanghai Economic Research* **2019,** 10, 1-17.
30. Song, Y., Study on the composition of family age members, public service and consumption of floating population. *Population and Development* **2019,** 25(2), 86-96.
31. Pan, Q.; Gao, W., Test and analysis on the effect of urbanization on residents' consumption in China. *Macroeconomic Research* **2014,** 1, 118-125.
32. Tan, J.; Yang, Y., The impact of population mobility and aging on rural residents' consumption. *Population Journal* **2012,** 6, 9-15.
33. Li, Z.; Yuan, Z., Consumption, investment, net export and economic growth in China. *Statistical Research* **2009,** 26(2), 39-42.
34. Zheng, X.; Zhang, Y., "Consumptive integration": the urban life practice and struggle of the new generation of migrant workers -- An empirical study based on the Yangtze River Delta urban agglomeration. *Journal of China University of Geosciences: Social Science Edition* **2016,** 16(1), 123-134.
35. Gao, B.; Wang, H., Empirical analysis of real estate price fluctuation and residents' consumption in Yangtze River Delta. *Industrial Economics Research* **2011,** 1, 1-10.
36. Borjas, G. J., The economics of immigration. *Journal of economic literature* **1994,** 32, (4), 1667-1717.
37. Liu, X.; Zhao, Y., The influence of labor mobility on urban labor market. *Economics (quarterly)* **2009,** 2, 693-710.
38. Peri, G., The effect of immigration on productivity: Evidence from US states. *Review of Economics and Statistics* **2012,** 94, (1), 348-358.
39. Wu, P., Squeeze downward or promote upward——Theoretical analysis of the relationship between urban unemployed and laid-off workers and migrant workers' employment. *Exploration and Free Views* **2007,** 4, 34-38.
40. Zhong, X., Labor mobility and wage differentials. *Chinese Social Sciences* **2006,** 1, 36-46.
41. Kou, Z.; Liu, X. *Report on city and industrial innovation in China*; Fudan Institute of Industrial

Development, Fudan University: 2017.

Chapter 6. Spatiotemporal dynamics of population mobility and mechanism analysis at street block scale: A case study in Qingdao, China

6.1 Introduction

The population mobility within the city greatly affects the urban vitality. At the same time, the urban vitality represented by the spatial and temporal distribution of urban population provides a new perspective for the study of population mobility within the city. Urban vitality was first described by Jacobs [1] as “liveliness and variety attract more liveliness; deadness and monotony repel life” in *The Death and Life of Great American Cities*. Then Jacobs further emphasized that urban vitality is highly influenced by the dense population related to social and economic activities, and human activities and their interaction greatly shape the attraction, diversity and prosperity of urban places [2, 3]. On the one hand, the vibrant urban space promotes the social exchange and interaction, ensuring the long-term sustainable development of the city [4, 5]; on the other hand, it improves the residents' subjective feelings of urban space, which is essential for the well-being and innovation of urban residents [6, 7]. Therefore, a comprehensive understanding of urban vitality is necessary for urban management and urban planning.

At present, urban vitality has attracted the attention of many disciplines, such as urban planning, geographic information science and social science. The research on urban vitality focuses on two aspects: portraying urban vitality and its determinants [5, 8-12]. Field research can directly investigate human activities, interactions and life experiences, so it is considered to be the most effective way to capture urban vitality [13]. However, the high cost, time-consuming and limited sample size further limit the use of this method. Subsequently, point of interest (POI) data [14], house price data [15] and land use data are regarded as representatives of human activities and interactions, and they are successfully adopted to indirectly represent urban vitality. Another research stream is devoted to reveal the determinants of vitality. Through the establishment of multiple regression and binomial regression, they found the close relationship between human activities and demographic with socio-economic factors [16, 17]. Urban built environment is also widely considered to have a significant impact on urban vitality [2, 13]. These findings provide valuable implications for the study of urban vitality. Nevertheless, the quantitative evaluation of the spatial dynamics of urban vitality remains a challenging issue, whether it is the definition of vitality itself or the mining of determinants.

In the era of big data, multisource urban big data are emerging, such as mobile phone data [5, 18, 19], global positioning data [20, 21], social media data [22-24], public transport smart card data [5] and Wi-Fi access point data [25]. These datasets contain abundant human activity and interactive information in the city [16, 26, 27]. They can capture a large number of human activities, rhythms and preferences at social perception level [28]. By generating multifaceted images of urban vitality, they bring deeper insights into the quantification of it. The urban vitality represented by human activities is gradually accepted by a wide range of scholars, which also conforms to Jacobs' definition of urban vitality, that is, the active street life created by the presence of pedestrians at any time of the day. Recently, De Nadai, et al. [29] derived a dynamic indicator from the density of mobile Internet records and explores the urban vitality of six Italian cities. Yue, Zhuang, Yeh, Xie, Ma and Li [17] quantified the vitality of Shenzhen based on the density of mobile users, and revealed the impact of mixed land use. Huang, et al. [30] combined multi-source urban big data to evaluated and characterized the urban vitality of Shanghai. These advanced studies further deepen the understanding of urban vitality.

However, there are still some limitations in the measurement of urban vitality and the exploration of influencing factors based on multisource urban big data. Firstly, the proxy of urban vitality used in previous studies did not consider the spatiotemporal dynamics of urban vitality, only using static datasets to describe the urban vitality at a certain time. As proposed by Montgomery [3], one of the most significant characteristics of urban vitality is the dynamic change in the spatial and temporal dimensions. Secondly, the research on urban vitality based on multisource urban big data is all concentrated in highly developed cities or high-density cities, such as Beijing, Shanghai and Shenzhen in China, and there is a lack of research on other types of cities in China. Different Chinese cities have different urban forms and are different from those in western countries, so it is necessary to explore whether the current theory of urban vitality is applicable to different types of Chinese cities [17, 31, 32]. Thirdly, many scholars have explored the influencing factors of urban vitality based on different analysis methods, including multivariate factor analysis, regression analysis and principal component analysis. However, these studies ignore the spatial effect of urban vitality, which will inevitably have an impact on the results, and further cause the misunderstandings of urban researchers and urban planners [18, 31, 33]. Based on the aforementioned studies, our work is in accordance with Jacobs's theory: a successful city street must exhibit different flows of people in different time periods.

As a new kind of data with accurate spatiotemporal information, Baidu heat map data has great potential and advantages in the study of human behavior and environment. For example, the data has a wide spatial coverage based on the number of mobile phone users accessed by computer terminals and mobile terminals, can provide a large number of samples and has a relatively low data acquisition cost. Therefore, it is widely used in reflecting regional characteristics, studying urban spatial relations and exploring residents' spatiotemporal relations, providing new perspectives, new methods and data sources for geographical research, urban planning and public participation [34]. In particular, the data has two outstanding advantages in the study of urban vitality: (1) the georeferenced Baidu heat map data can indicate well-timed population distribution on the temporal scale (for example, by hour, by weekdays and weekends) and the spatial scale (for example, traffic analysis zone, street block and grid scale); (2) cover the entire user group. Therefore, these data can be used to reflect the spatial and temporal changes of urban vitality. In general, previous studies did not take into account diurnal variations in activity, including differences between weekdays and weekends and between daytime and nighttime. Therefore, this study attempts to use Baidu heat map data in each day (weekdays and weekends) as a proxy for urban vitality to investigate the Changes of urban population mobility characteristics in different time periods.

The urban built environment plays an important role in creating and maintaining urban vitality and affecting urban population mobility[35, 36]. Although previous studies have emphasized the profound impact of built environment indicators such as land use on urban vitality and population mobility, the debate on built environment indicators such as building density and accessibility on urban vitality has never ceased. On the one hand, Jacobs [36] pointed out that there is a dialectical linkage between high-density environment and maintaining urban vitality (represented by population mobility). Contemporary economists Glaeser [37] and Storper [38] also expounded the relationship between spatial characteristics and the “triumph of cities”. On the other hand, starting from the cognition of western industrial cities, Wirth [39] and Simmel [40] emphasized that the high density of population leads to the negative attitude, hostility and ignorance of residents. They believed that overcrowding, high density and diversity will cause psychological pressure and negative health impact. In fact, the current debate has two limitations. First of all, most of the empirical evidence is based on western developed cities and high-density cities. Some cities in China are becoming emerging new urbanization centers. New research is needed on these cities at different development stages to further test relevant arguments. Second, most of Jacobs’ views on urbanization are based on her anecdotal observations. At present, such observations can be quantitatively studied by using the newly acquired “urban big data”. Therefore, based on the 5Ds principle proposed by Belzer and Autler [41], this paper established a comprehensive measurement system of urban built environment from five aspects of density, diversity, design, distance to transit and destination accessibility.

This research makes three primary contributions. Firstly, the potential data source (Baidu heat map data) was taken as the proxy of urban vitality to explore urban population mobility for the first time and its applicability was verified, which reflects the multifaceted nature of urban big data to represent urban vitality and population. Secondly, a complete and diverse framework was constructed for the urban built environment, including density, diversity, design, distance to transit and destination accessibility. Taking Qingdao, China as the study area, which enriches the research of different types of cities. Thirdly, this study distinguished urban vitality of weekdays, weekends, daytime and nighttime, thereby revealed the different lifestyles of the residents in different time periods. With the help of the spatial econometric model, the relationship between the urban built environment and urban vitality was discussed under the consideration of the spatial effect of urban vitality. The results can not only expand the areas related to urban vitality and urban mobility, but also provide meaningful guidance for urban design and urban planning.

The reminder of this chapter is organized as follows. Section 6.2 introduces research area and data. Section 6.3 outlines the methodology. Section 6.4 focus on the results and discussion. Scetion 6.5 makes the summary of this chapter.

6.2 Study areas and data

6.2.1 Study areas

This chapter makes an exploratory case study in Qingdao, China (Fig. 6-1). Due to the special background and development process, Qingdao occupies an important position in the history of modern urban planning in China. First of all, Qingdao is a new city built under the guidance of western urban planning theory and urban development concepts. After being occupied by Germany in 1897, Qingdao completed the transformation from a small fishing village to a modern city. The growth of population and economy laid a prominent position in the urban system of China. Although Qingdao has experienced several urban planning in the following hundred years, its urban construction is still deeply influenced by western urban planning thoughts. Secondly, special geographical conditions have shaped Qingdao's unique urban form. Under the constraints of mountains and the sea, different urban design methods have been adopted in different regions. For example, only in the old southern urban area occupied by Germany, there are different street forms such as network format, free style and magnificent style. This creates an opportunity for us to study the practice of western theories related to urban vitality in Chinese cities and the interactive relationship with urban built environment. In this paper, we have selected the southern coastal area (the main urban area) that best reflects western urban planning concepts as the research area, namely Shinan district, Shibei district, Licang district and Laoshan district. Previous quantitative studies of urban vitality were conducted at different scales, from the municipal level [3] to the district level [17, 29]. We adopted a grid with a 200m*200m spatial resolution [42] to provide spatial dynamics analysis of urban vitality at the street block level.

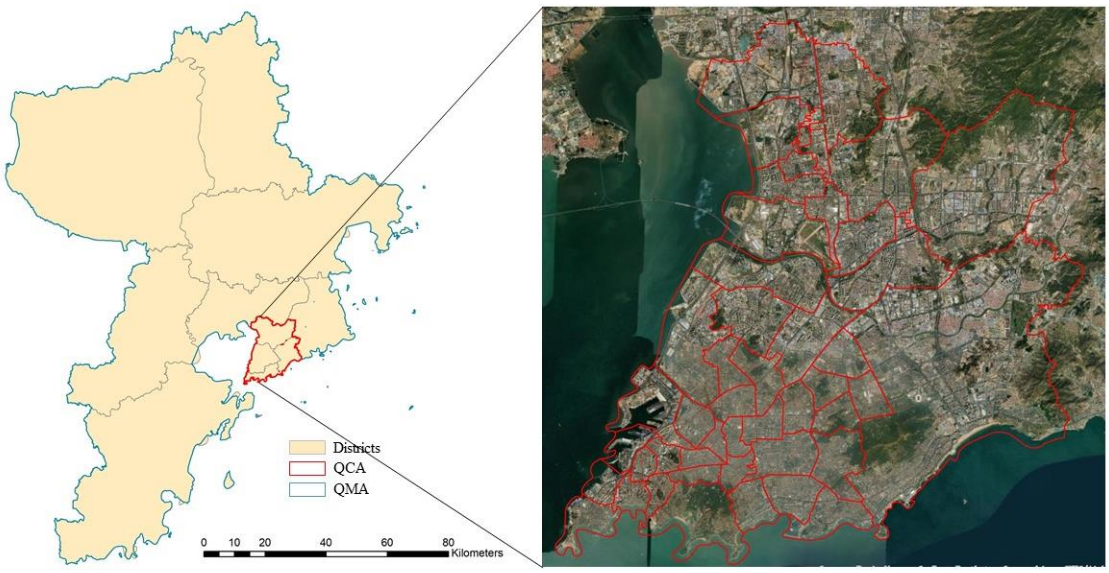


Fig. 6-1. Study area: Qingdao, China (QCA: Qingdao central area; QMA: Qingdao municipal area).

6.2.2 Research data

The data used in this study included Baidu heat map data as the proxy of urban vitality and urban spatial information data, such as point of interests (POIs) for land use function, road networks and building footprint information.

Baidu is the largest search engine and website in China. Its terminal market share accounted for 73.5%, covering 97.5% of all Chinese users, and the average daily response volume reached 6 billion times. In 2011, Baidu launched big data visualization products, namely Baidu heat map. Baidu heat map data is based on the location information obtained by mobile phone users when they access Baidu products (Baidu search, Baidu weather, Baidu map, and Baidu music, etc.), comprehensively calculate the heat value of population flow in different regions, and accurately reflect the degree of population aggregation in the region. Baidu heat map is updated every 15 minutes to obtain the dynamic information of population distribution in real time.

As a big data application with hundreds of millions of users, Baidu heat map has a wide range of coverage, high-resolution spatiotemporal characteristics and easy access, which embodies the potential and value in urban research. At present, the urban spatial analysis based on this data has been widely carried out, including urban spatial structure, urban job-housing balance, polycentric urban development, and urban population aggregation [43-46]. In this chapter, the processed Baidu heat map data was used as the proxy of urban vitality.

Previous studies have shown that urban population has similar evolutionary patterns from Monday to Friday (weekdays), Saturday and Sunday (weekends). Therefore, this chapter chooses December 14 (weekday) and 15 (weekend) in 2018 as the research time, calls the website application programming interface (API) through the Python programming language, obtains data at 1-hour intervals, and finally acquires a total of 36 Baidu heat maps (7:00-24:00). The population distribution data of spatial grid units generated after image processing is stored in the database.

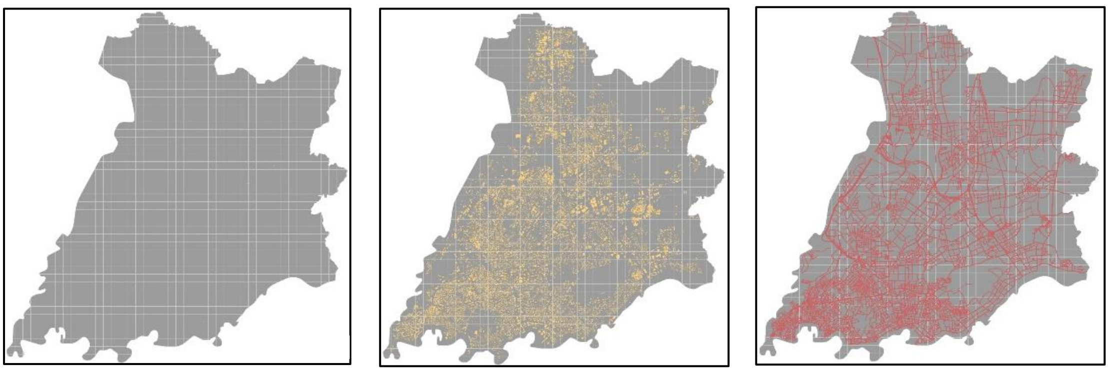
The urban spatial information data used in this chapter includes point of interest (POI) data, building footprint data and road network data. In China, the POI categories are in accord with land-use classifications. POI data not only has more flexibility in research scale, but also has much finer statistical granularity, and can also be used to express the preferences of population and urban social functions. Therefore, POI data are thus used to reflect land use. The POI data are obtained from AutoNavi using API, which is a network map platform and a location-based service provider in China. A total of 69980 POIs are obtained, and this paper classifies these points into 7 types according to the Code for Classification of Urban Land Use and Planning Standards of Development Land (GB50137-2011). The details of the POIs in each category are shown in Table 6-1.

Table 6-1. Categories of POIs.

POI data category	POI data type	Primary information
Traffic-related POI (TPOI)	Transport hub	Railway station, Airport, Port, etc.
	Traffic facilities	Bus station, Subway station, etc.
Public-related POI (PUPOI)	Science, Education, Cultural and Sports service	School, Academy, Museum, Gym, etc.

	Medical care	Hospital, Clinic, Drugstore, etc.
	Governmental agencies, Social organization	Governmental agencies, Neighborhood committee, etc.
Park-related POI (PAPOI)	Places of interest	Cultural tourism attraction, Natural reserve, etc.
	Urban park	Park, Square, Playground, etc.
Industrial-related POI (IPOI)	Factory	Cement plant, Steelworks, Quarry, etc.
	Industrial building	Industrial building
Housing-related POI (HPOI)	Residential community	Residential community
	Accommodation service	Hotel, Guesthouse, etc.
Consumption-related POI (CPOI)	Shopping service	Shopping mall, Department store, Supermarket, etc.
	Catering service	Restaurant, Pub, Café, Dessert shop, etc.
	Life service	Barbershop, Dry cleaner, Telecom office, etc.
	Recreation and entertainment	Cinema, KTV, Game room, etc.
Business-related POI (BPOI)	Corporate business	Technology company, Consulting company, Design company, etc.

The building footprint data is also obtained from AutoNavi using API, including all building space coordinates, basic contours and floor information in the study area of 2018, which is essential for measuring physical urban spaces. The urban road network data is crawled from the Open Street Map (OSM) of 2018. It is believed that the data quality of the OSM for major cities in China is good, and this dataset has been used to measure the built environment. After processing, the road network data we obtained includes primary and secondary arterial roads and branches. The details of the urban spatial information data are shown in Fig. 7-2. Table 7-2 displays the basic descriptive statistics of the variables.



(a) Analysis grid in the central area of Qingdao

(b) Building footprints in the central area of Qingdao

(c) Road network in the central area of Qingdao

Fig. 6-2. (a) Analysis grid in the central area of Qingdao. (b) Building footprints in the central area of Qingdao. (c) Road network in the central area of Qingdao.

Table 6-2. Descriptive statistics for the variables in grid.

Variables	Mean	S.D.
Density		
Point of interest (POI) density	234.4	364.3
Diversity		
Mixed function degree	0.7	0.3
Proportion of traffic-related POI	17	29
Proportion of housing-related POI	33	75
Proportion of public-related POI	64	105
Proportion of business-related POI	55	110
Proportion of park-related POI	2	8
Proportion of consumption-related POI	62	121
Proportion of industrial-related POI	1	3
Design		
Floor area ratio (FAR)	2.5	2.4
Building coverage ratio (BCR)	0.5	0.4
Distance to transit		
Distance to bus station and subway station	1.4	1.2
Distance to central business district	3.6	2.1
Destination accessibility		
Road density (RD)	24.3	22.4
Road intersection density (RID)	191.2	282.1

6.3 Methods

In this chapter, we firstly investigated the spatiotemporal dynamics of urban vitality represented by Baidu heat map using exploratory spatial data analysis and then examined the linkage between the urban vitality and urban built environment using spatial econometric model. Fig. 7-3 is the research flow in this chapter. Meanwhile, this chapter quantifies the urban built environment indicators from five aspects, namely density, diversity, design, distance to transit and destination accessibility. Table 6-3 lists the detailed information of all variables.

Table 6-3. List of variables used in analysis.

Main type	Name	Description
Density	Point of interest (POI) density	Point of interest (POI) count of each analysis grid
Diversity	Mixed function degree	Mixed function degree based on the POI data and information entropy
	Proportion of traffic-related POI	Proportion of traffic-related POI of each analysis grid
	Proportion of housing-related POI	Proportion of housing-related POI of each analysis grid
	Proportion of public-related POI	Proportion of public-related POI of each analysis grid
	Proportion of business-related POI	Proportion of business-related POI of each analysis grid
	Proportion of park-related POI	Proportion of park-related POI of each analysis grid
	Proportion of consumption-related POI	Proportion of consumption-related POI of each analysis grid
	Proportion of industrial-related POI	Proportion of industrial-related POI of each analysis grid
Design	Floor area ratio (FAR)	The ratio of gross floor area to building footprint of each analysis grid
	Building coverage ratio (BCR)	Building density within each analysis grid
Distance to transit	Distance to bus station and subway station	The average air distance to the closest bus station and subway station
	Distance to central business district	The average air distance to the closest central business district
Desitination accessibility	Road density (RD)	The ratio of total length of road in each analysis grid
	Road intersection density (RID)	Road intercetion count of each analysis grid

Tan, Huang, Zhao, Yu, Leng and Feng [44] put forward a method for forecasting regional

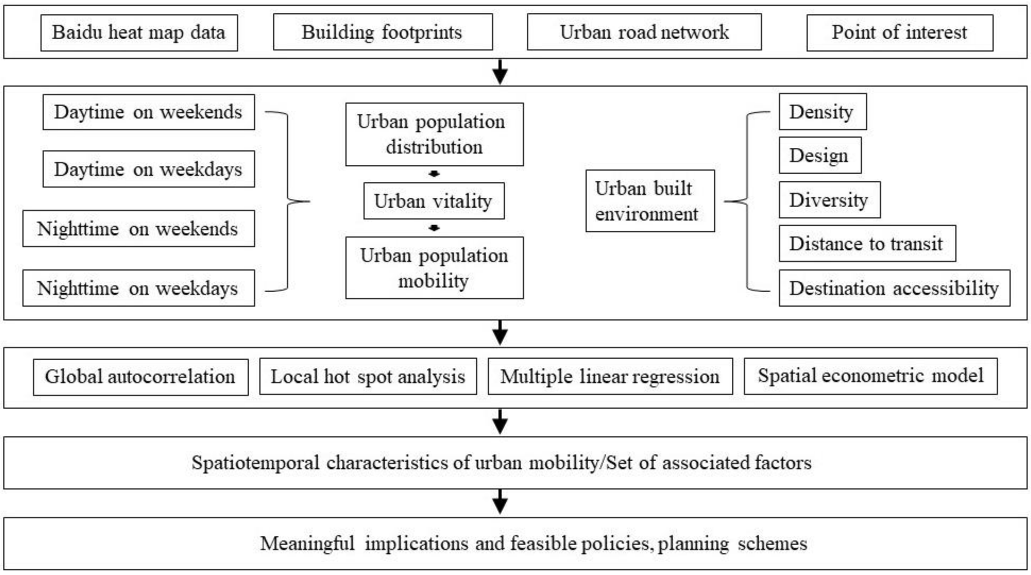


Fig. 6-3 Research flow chart of this chapter.

population based on Baidu heat map data. After the improvement, a method to predict regional population density is adopted:

$$D = \frac{\sum_{i=1}^{n} d_i * m_i * s}{A} \qquad (1)$$

where D represents the population density within the grid analysis unit (500m*500m) at a specific time, di represents the population density of color i, mi donates the number f pixels of color i, s is the pixel area of a single pixel, m represents the type of color and A is the grid analysis unit (The population density represented by different colors is based on Baidu's official legends. Among them, the red indicates a population density of 60/hm2, the orange of 40–60/hm^2, the yellow of 20–40/hm^2, the pale green of 10–20/ hm^2, and the deep green of<10/hm^2.).

In order to avoid population preferences (age, background, etc.) having a greater impact on the predicted population density, we further deal with formula (1) as follows:

$$P = \frac{D_{ta}}{\sum D_{ta}} \qquad (2)$$

where P is the urban vitality, t represents the time, a donates the grid analysis unit, Dta is the population density of the analysis unit at the time t, $\sum$Dta is the population density of all analysis units at the time t.

In this chapter, we divide urban vitality into four types, namely, urban vitality during the daytime on weekdays (UVDW), urban vitality during the nighttime on weekdays (UVNW), urban vitality during the daytime on weekends (UVDE) and urban vitality during the nighttime on weekends (UVNE). Among them, the UVDW and UVDE are represented by the average value of the vitality from 9:00 to 17:00, the UVNW and UVNE are represented by the average value of the vitality from 18:00 to 24:00.

6.4 Results

6.4.1 Spatiotemporal patterns of urban vitality

Spatiotemporal patterns of urban vitality in Qingdao, as indicated by Baidu heat map data, are displayed in Fig. 6-4, 6-5, 6-6 and 6-7. Although all these patterns show certain similarities, the urban vitality of different regions varies significantly over time. During the daytime on weekdays, the urban vitality center appears on Zhongshan Road, the Municipal Government, Taitung and Licang business district, where Zhongshan Road, close to Qingdao railway station and the southern coastal tourism area, attracts a large number of people. As comprehensive regional centers, the Municipal Government, Taitung and Licang business district also have great potential in attracting people flow. As a result, urban activity is highest in these areas. Compared with the urban vitality on weekdays' daytime, the urban vitality on weekdays' nighttime has obvious expansion phenomenon, Taking the above areas as the center, the urban vitality in the peripheral areas has increased significantly. The urban vitality shows the spatiotemporal difference characteristics of "contraction in the daytime and expansion in the nighttime". On weekends, the spatiotemporal change of urban vitality is the opposite of that on weekdays, showing the changing characteristics of "expanding in the daytime and contracting at nighttime". The time difference of urban residents' activity purposes is one of the factors that cause the spatiotemporal change of urban vitality. The traditional life style and daily routines rules govern the types of activities of urban residents, and the corresponding characteristics between time and content are obvious. During weekdays, residents generally go through the process of "employment-leisure-rest", and during weekends, they generally go through the process of "rest -leisure". It can be seen that the spatial-temporal change of urban

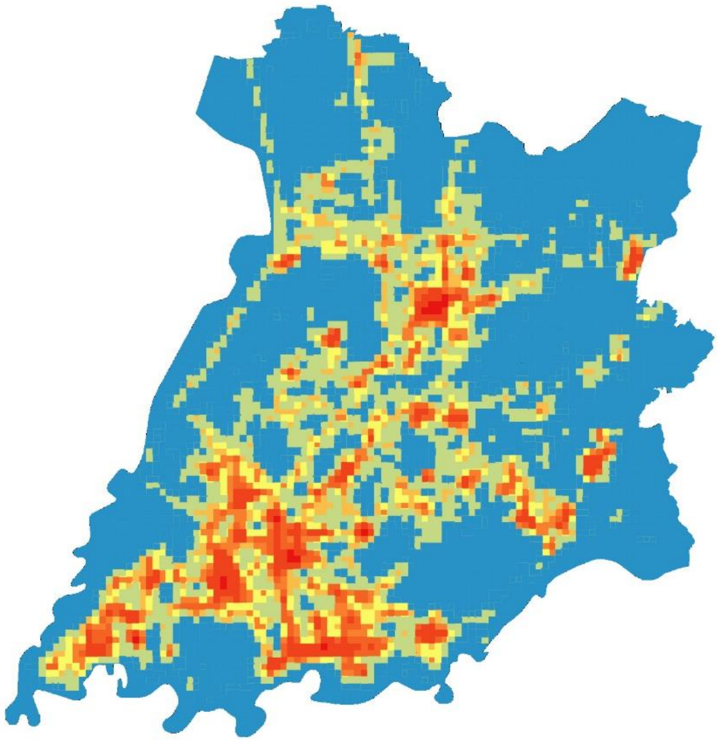


Fig. 6-4 Urban vitality in Qingdao central area (during daytime on weekdays).

vitality is related to the difference of residents' activities.

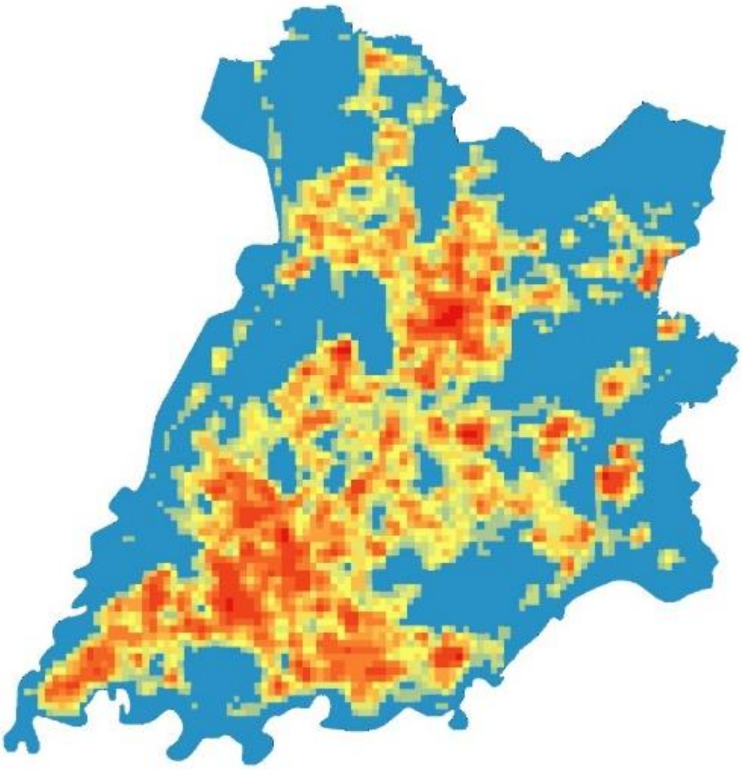


Fig. 6-5. Urban vitality in Qingdao central area (during nighttime on weekdays).

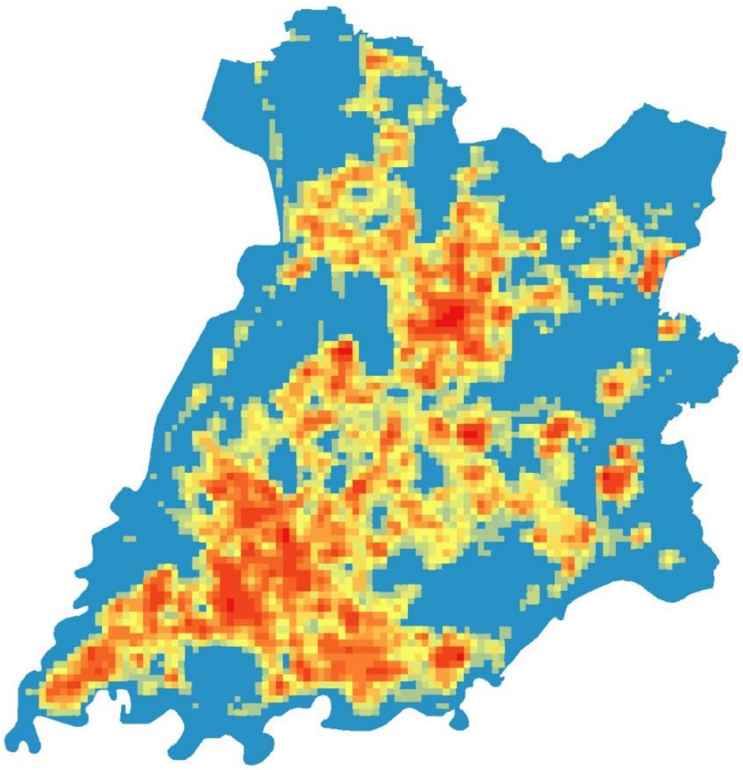


Fig. 6-6. Urban vitality in Qingdao central area (during daytime on weekends).

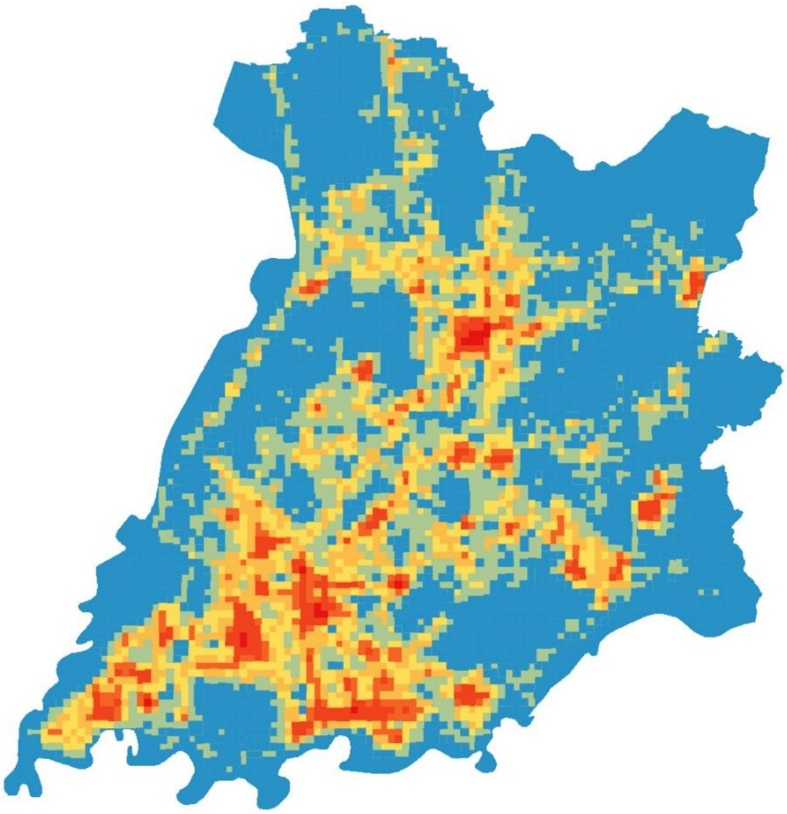


Fig. 6-7 Urban vitality in Qingdao central area (during nighttime on weekends).

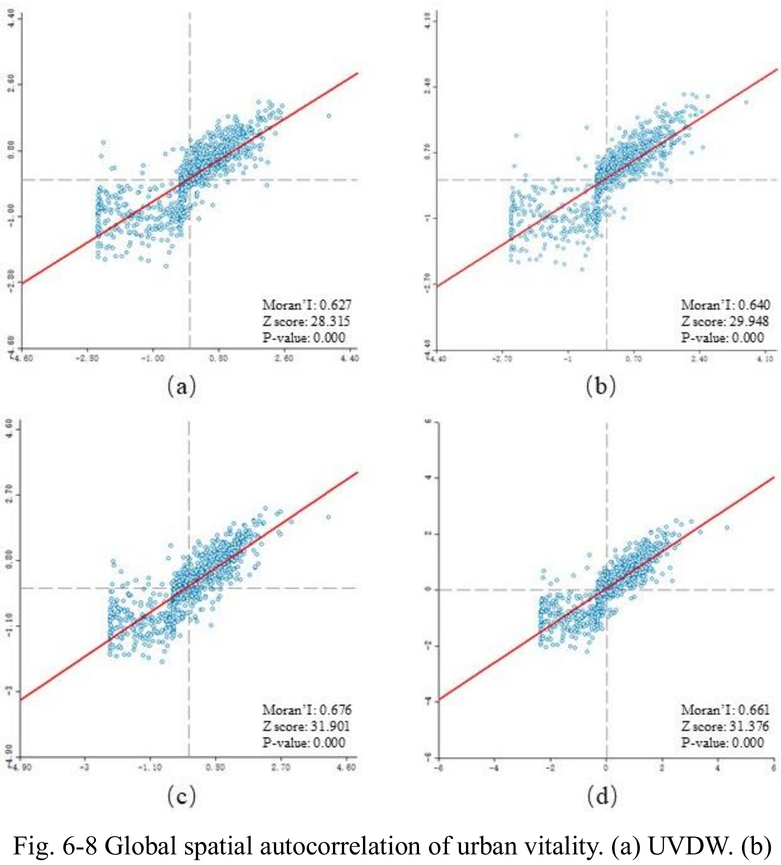


Fig. 6-8 Global spatial autocorrelation of urban vitality. (a) UVDW. (b) UVNW. (c) UVDE. (d) UVNE.

The results further demonstrate the polycentric structure of vitality, which in line with the urban growth of Qingdao. No matter at any time, Zhongshan Road area, Municipal Government area, Licang business district, Taitung are the vitality center in the main urban area. The Moran test of urban vitality is shown in Fig. 7-8, and spatial autocorrelation can be observed for all Z values greater than 20. On the one hand, it confirms the importance of spatial patterns of neighborhood vitality in Fig. 6-4, 6-5, 6-6 and 6-7, on the other hand, it emphasizes the necessity of considering variable spatial effect in regression model. Therefore, in the following chapters, with the help of the spatial econometric model, we focus on another factor that affects the urban vitality, namely urban built environment.

7.4.2 Impacts of urban built environment on urban vitality

After the multicollinearity test, the VIF values of all variables were less than 7.5, so a total of 15 variables were put into the regression model. Table 6-4 exhibits the results of the OLS regression and spatial econometric model. The fitting coefficients of SAR (0.752) and SAC (0.779) are better than OLS (0.528), while the value of Log Likelihood is higher than OLS and the results of SAR and SAC judged by AIC and SC are less than OLS. These pieces of evidence indicate that the spatial econometric model is better than the OLS regression due to the consideration of spatial effects between variables. By comparing Lagrange Multiplier (LM) and Robust LM, the SAC is superior to the SAR. Therefore, SAC was used to examine the impact of urban built environment on urban vitality.

Table 6-4. Results of the comparison of OLS regression and spatial econometric model.

	OLS	**SAR**	**SAC**
R^2	0.528	0.752	0.779
Log likelihood	-1227.07	-1079.29	-1045.97
Akaike info criterion	2460.14	2164.60	2099.94
Schwarz criterion	2475.35	2179.81	2120.23
Lagrange multiplier (LM)	-	159.08	219.22
Robust LM	-	0.61	60.75***

The results show that the urban built environment quantified by five aspects of design, density, diversity, distance to transit and destination accessibility has a significant impact on urban vitality, but the degree of impact varies in different periods (Table 7-5 and Table 7-6). Among them, urban functional density and central business district distance have the strongest correlation with urban vitality, followed by building floor area ratio, road density and Road intersection density. Mix land use has a weak positive correlation with urban vitality.

Table 6-5. SAC model for the relationship between urban vitality on weekdays and urban built environment.

Variable	Regression 1			Regression 2		
	Coefficient	T-statistic	Std. Error	Coefficient	T-statistic	Std. Error
(Constant)	-0.032	-1.164	0.027	-0.033	-1.159	0.029
Design						
Building coverage ratio	-0.032*	-1.164	0.027	-0.033*	-1.159	0.029
Floor area ratio	-0.032***	-1.164	0.027	-0.033***	-1.159	0.029
Density						
Density of POI	0.257***	9.314	0.027	0.131***	7.926	0.016
Diversity						
POI function mix	0.068***	2.638	0.026	0.051***	3.566	0.014
Proportion of traffic-related POI	0.066***	4.566	0.014	0.064***	4.272	0.015
Proportion of public-related POI	0.113***	6.865	0.016	0.024*	1.843	0.013
Proportion of park-related POI	0.066***	4.364	0.015	0.074***	5.896	0.015
Proportion of industrial-related POI	0.027	0.102	0.013	0.155	1.003	0.017
Proportion of housing-related POI	0.086***	5.412	0.016	0.096***	5.869	0.018
Proportion of consumption-related POI	0.129***	7.189	0.018	0.117***	2.598	0.037
Proportion of business-related POI	0.144***	7.871	0.018	0.105***	2.761	0.038
Destination accessibility						
Road density	0.156***	5.398	0.028	0.137***	4.753	0.028
Road intersection density	0.103***	3.221	0.018	0.096***	2.501	0.017
Distance to transit						
Distance to bus station and subway station	-0.068*	-1.905	0.036	-0.075**	-2.008	0.037
Distance to central business district	-0.309***	-6.743	0.045	-0.345***	-7.218	0.047
Dependent variable		UVDW			UVNW	
R^2		0.779			0.782	

Note: *p<0.10, **p<0.05, ***p<0.01.

Table 6-6. SAC model for the relationship between urban vitality on weekends and urban built environment.

Variable	Regression 3			Regression 4		
	Coefficient	T-	Std.	Coefficient	T-statistic	Std. Error

		statistic	Error			
(Constant)	-0.025	-1.046	0.023	-0.006	-2.201	0.014
Design						
Building coverage ratio	0.108^{*}	1.202	0.024	0.063^{*}	1.102	0.026
Floor area ratio	0.228^{***}	9.354	0.012	0.291^{***}	11.542	0.025
Density						
Density of POI	0.103^{***}	7.026	0.013	0.147^{***}	8.372	0.017
Diversity						
POI function mix	0.048^{***}	3.727	0.014	0.044^{***}	2.816	0.015
Proportion of traffic-related POI	0.053^{***}	3.961	0.011	0.044^{***}	2.902	0.015
Proportion of public-related POI	0.015	0.089	0.014	0.027^{*}	1.841	0.014
Proportion of park-related POI	0.097^{***}	5.549	0.016	0.103^{***}	5.014	0.016
Proportion of industrial-related POI	0.032	8.253	0.016	0.176	0.698	0.019
Proportion of housing-related POI	0.107^{***}	5.385	0.032	0.127^{***}	7.234	0.017
Proportion of consumption-related POI	0.068^{**}	2.108	0.033	0.097^{***}	3.082	0.031
Proportion of business-related POI	0.131^{***}	3.891	0.025	0.123^{***}	3.402	0.037
Destination accessibility						
Road density	0.164^{***}	4.421	0.017	0.091^{***}	3.361	0.026
Road intersection density	0.708^{***}	3.226	0.031	0.599^{***}	2.757	0.018
Distance to transit						
Distance to bus station and subway station	-0.066^{**}	-2.124	0.039	0.002	0.121	0.021
Distance to central business district	-0.317^{***}	-8.003	0.038	-0.260^{***}	-10.168	0.025
Dependent variable		UVDE			UVNE	
R^2		0.826			0.748	

Note: *p<0.10, **p<0.05, ***p<0.01.

Density. Density is closely related to the maintenance of urban vitality, which is consistent with the analysis results of Long and Huang [31]. POI density reflects the distribution of convenience amenities in urban streets, covering various aspects such as catering, shopping, science and education, public services, medical services, life services, and government agencies. As the main place for production, circulation and consumption of material products, the denser the urban functional amenities, the more able to meet the needs of urban residents, thereby promoting social and economic activities and enhancing regional vitality. The concentration of urban population further promotes urban vitality by making the urban environment safer and encouraging face-to-face interaction between people.

Design. The urban form also has a positive influence on shaping the urban vitality. After the reform and opening up, China's urban space has undergone extensive development in the vertical dimension. More than half of the world's skyscrapers are located in China. The evidence in Qingdao is consistent with the views of Huang, Zhou, Li, Song, Cai and Tu [30], Ye, Li and Liu [12] and Tu, Zhu, Xia, Zhou, Lai, Jiang and Li [10], that the increase in building height has a positive contribution to the urban vitality. On the one hand, the building height not only determines the urban form, but also directly affects the capacity of urban space to carry all aspects of human activities. On the other hand, the vertical structure of urban buildings has a psychological impact on people's desire to stay [47]. However, based on the evidence in Qingdao, we also found that the results of studies different from Huang, Zhou, Li, Song, Cai and Tu [30] and Ye, Li and Liu [12] are that the effect of increasing the horizontal density of buildings on urban vitality is almost negligible compared with increasing the height of buildings. This shows that in the future, when considering the cultivation of urban vitality in urban spatial planning and design, it should focus on filling vertical space, rather than filling horizontal space to enhance the vitality of high-density and highly developed cities such as Beijing, Shanghai and Shenzhen. This also highlights the necessary to study Chinese cities in different stages of development.

Diversity. In theory, diversity is considered to be the main source of urban vitality. However, some research reported conflicting results. For example, De Nadai, Staiano, Larcher, Sebe, Quercia and Lepri [29] found no significant correlation between the vitality of Italian cities and the diversity of urban functions. Our research supplements the evidence of Chinese cities and contributes to the related controversy. The results based on Qingdao indicate that the mixing of urban functions can positively affect the urban vitality, whether it is on weekdays, weekends, daytime or nighttime. This is because the diversity of urban functions encourages people to stay and interact in urban spaces, thereby increasing the vitality of the region. However, the impact of urban functional diversity is limited, and this result is consistent with the findings of Tu, Zhu, Xia, Zhou, Lai, Jiang and Li [10]. The urban vitality cultivation method considering functional diversity does not naturally improve the urban vitality in Qingdao. At the same time, we also found differences in the impact of different categories of POI on urban vitality. For example, increasing the proportion of consumption-related POI and business-related POI can significantly improve the urban vitality on weekdays, and increasing the proportion of housing-related POI and park-related POI will have a greater impact on the urban vitality on rest days. This discovery can assist urban planning and regional design in the future, and actively cultivate the neighborhood vitality.

Destination accessibility. Good accessibility can promote urban vitality. Road density and road intersection density are direct reflections of accessibility. Evidence from Qingdao indicates that both have a significant positive impact on urban vitality. The size of the street blocks slows down the pace of life of urban residents, so that they have enough time to engage in social and consumer activities. The urban form of dense road and small blocks increases the possibility of random contact among urban residents. On the one hand, these blocks can avoid lengthy and tedious walks and create a vibrant urban space. On the other hand, the high-density road provides multiple path options and attracts various small community shops. Our conclusion on Qingdao is consistent with that of Long and Huang [31] and Huang, Zhou, Li, Song, Cai and Tu [30]. Therefore, the community form of dense road not only creates more opportunities for people to contact, but also improves the effective combination of land use, thereby making the city full of vitality. In particular, Qingdao’s Zhongshan road and Taitung, as blocks built under the guidance of western urban planning concepts, their small squares and dense road make it as the vibrant center of Qingdao from last century to the present.

Distance to transit. Distance to transit has a negative correlation with urban vitality. Among them, the distance to CBD has a significant negative impact on urban vitality, which indicates that with the increase of distance, urban vitality is declining. This is consistent with the analysis results of Wu, Ye, Ren and Du [11] and Long and Huang [31]. The city’s CBD highly integrates the region's economy, culture and technology, and has multiple functions such as finance, trade, service, exhibition and consulting, which provides great convenience for the social life of urban residents. Therefore, the closer to the CBD, the closer the interaction between people, and the stronger the neighborhood vitality. At the same time, we also found that the distance to urban public transportation stations also has a negative correlation with urban vitality, which shows that a perfect public transportation system can improve accessibility, thereby cultivating vitality in the future. Convenient travel ways are an important issue in attracting human activities and neighborhood interaction. High density and wide coverage of the ways to travel is conducive to convenient travel in the city, so as to cultivate vitality. At present, Qingdao’s subway system has only opened three lines. In the future, more lines will cover the suburbs of Qingdao, making it more convenient to connect the regional center and the main urban area, which is crucial to cultivate local vitality.

6.5 Summary

This chapter attempts to use Baidu heat map data as a proxy for urban vitality, depicting the spatial and temporal dynamics of Qingdao's urban vitality, focusing on analyzing the impact of the urban built environment on urban vitality. This study combines the Baidu heat map data and exploratory spatial data analysis method, complements the multifaceted measurement of urban vitality, and makes contributions to the existing literature. The application of spatial econometric model improves the reliability of the relationship between urban built environment and urban vitality. In addition, it surpassed the earlier research, and built a comprehensive measurement system for the urban built environment through five aspects: density, design, diversity, distance to transit and destination accessibility.

The study found that Qingdao's urban vitality showed certain differences over time. On the weekdays, the urban vitality shows the spatiotemporal characteristics of "daytime contraction, nighttime expansion", while on the weekends, it is the opposite. At the same time, the urban vitality shows a polycentricity, forming a spatial structure with Zhongshan road, Eastern business district, Licang business district and Taitung as regional centers, radiating to the surroundings. This spatial structure is consistent with the empirical research of Huang, Zhou, Li, Song, Cai and Tu [30] in Shanghai and the exploration of Tu, Zhu, Xia, Zhou, Lai, Jiang and Li [10] in Shenzhen.

Other findings in this paper are that the density of urban functions and the accessibility of destinations have a significant positive effect on urban vitality. For example, city managers should actively encourage real estate developers to come up with plans for small blocks, dense road networks, and wide coverage of convenient amenities. Although considering the form of urban buildings can also enhance the urban vitality. However, compared with other developed cities in China, such as Beijing, Shanghai and Shenzhen, increasing the vertical space of buildings can significantly improve the neighborhood vitality in Qingdao, while the expansion of the horizontal space of buildings is not obvious, which highlights the necessity of studying cities in different development stages in China. Diversity can also have a positive impact on urban vitality, this effect is also not obvious, which should be noted in planning practice. The CBD and allocation of public transport system also has a positive impact on urban vitality. Our results contribute to the debate and future planning practice of different types of urban vitality spaces.

The main contributions of this paper are as follows: (1) From the perspective of social perception, it verifies the possibility of Baidu heat map data as a proxy for urban vitality, and complements the data base and method of depicting urban vitality image. (2) Taking Qingdao as an example, a complete and systematic measurement of urban built environment is constructed and verified the influence on urban vitality at street block level. The proposed framework can also be applied to other cities in China. (3) The impact of relevant factors on urban vitality can be used to assess neighborhood vitality and determine appropriate design interventions in an evidence-based manner.

Although we believe that the research results of this paper are relatively reliable and important to understand the vitality of urban design, we think that there is still greater exploration space in some aspects. First of all, the single dataset portrays the urban vitality, which inevitably has deviations and lacks multiple perspectives. Future research may consider combining social media check-in data and mobile phone records data to build a complete, cross validated urban vitality proxy framework. Secondly, the definition of block scale should be more accurate, rather than the 200m * 200m grid proposed in some paper, which will provide more accurate insights, such as regulatory planning management unit (RUMP) and traffic analysis zones (TAZs). In a word, this paper verifies the possibility of Baidu heat map data as a proxy for urban vitality and explores the impact of urban built environment. On the one hand, it makes a good supplement for the relevant debate and measurement data source of urban vitality, on the other hand, it provides a certain data reference for urban planning practice.

Reference

1. Jacobs, J., *The death and life of great American cities*. Vintage: 2016.
2. Chhetri, P.; Stimson, R. J.; Western, J., Modelling the factors of neighbourhood attractiveness reflected in residential location decision choices. *Studies in Regional Science* **2006,** 36, (2), 393-417.
3. Montgomery, J., Making a city: Urbanity, vitality and urban design. *Journal of urban design* **1998,** 3, (1), 93-116.
4. Mayo, J. M., Naked city: The death and life of authentic urban places. In SAGE Publications Sage CA: Los Angeles, CA: 2011.
5. Sulis, P.; Manley, E.; Zhong, C.; Batty, M., Using mobility data as proxy for measuring urban vitality. *Journal of Spatial Information Science* **2018,** 2018, (16), 137-162.
6. Pinquart, M.; Sörensen, S., Influences of socioeconomic status, social network, and competence on subjective well-being in later life: a meta-analysis. *Psychology and aging* **2000,** 15, (2), 187.
7. Xu, Y.; Belyi, A.; Bojic, I.; Ratti, C., How friends share urban space: An exploratory spatiotemporal analysis using mobile phone data. *Transactions in GIS* **2017,** 21, (3), 468-487.
8. Lu, S.; Shi, C.; Yang, X., Impacts of built environment on urban vitality: regression analyses of Beijing and Chengdu, China. *International journal of environmental research and public health* **2019,** 16, (23), 4592.
9. Tang, L.; Lin, Y.; Li, S.; Li, S.; Li, J.; Ren, F.; Wu, C., Exploring the Influence of Urban Form on Urban Vibrancy in Shenzhen Based on Mobile Phone Data. *Sustainability* **2018,** 10, (12), 4565.
10. Tu, W.; Zhu, T.; Xia, J.; Zhou, Y.; Lai, Y.; Jiang, J.; Li, Q., Portraying the spatial dynamics of urban vibrancy using multisource urban big data. *Computers, Environment and Urban Systems* **2020,** 80, 101428.
11. Wu, C.; Ye, X.; Ren, F.; Du, Q., Check-in behaviour and spatio-temporal vibrancy: An exploratory analysis in Shenzhen, China. *Cities* **2018,** 77, 104-116.
12. Ye, Y.; Li, D.; Liu, X., How block density and typology affect urban vitality: An exploratory analysis in Shenzhen, China. *Urban Geography* **2018,** 39, (4), 631-652.
13. Azmi, D. I.; Karim, H. A., Implications of walkability towards promoting sustainable urban neighbourhood. *Procedia-Social and Behavioral Sciences* **2012,** 50, 204-213.
14. Joosten, V.; Van Nes, A. In *How block types influences the natural movement economic process: Micro-spatial conditions on the dispersal of shops and Café in Berlin*, Proceedings of the 5th International Space Syntax Symposium, TU Delft, Delft, The Netherlands, 2005; 2005.
15. Nicodemus, A. G., Fuzzy vibrancy: Creative placemaking as ascendant US cultural policy. *Cultural Trends* **2013,** 22, (3-4), 213-222.
16. Xu, Y.; Belyi, A.; Bojic, I.; Ratti, C., Human mobility and socioeconomic status: Analysis of Singapore and Boston. *Computers, Environment and Urban Systems* **2018,** 72, 51-67.
17. Yue, Y.; Zhuang, Y.; Yeh, A. G.; Xie, J.-Y.; Ma, C.-L.; Li, Q.-Q., Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy. *International Journal of Geographical Information Science* **2017,** 31, (4), 658-675.
18. Jiang, J.; Li, Q.; Tu, W.; Shaw, S.-L.; Yue, Y., A simple and direct method to analyse the influences of sampling fractions on modelling intra-city human mobility. *International Journal of Geographical Information Science* **2019,** 33, (3), 618-644.
19. Tu, W.; Zhong, S. In *Simulation and Experiment of Terahertz Wave Detection of Marine*

Protective Coating, FZU-OPU-NTOU joint symposium on Advanced Mechanical Science & Technology for Industrial Revolution 4.0, 2016; Springer: 2016; pp 141-150.

20. Tu, W.; Cao, R.; Yue, Y.; Zhou, B.; Li, Q.; Li, Q., Spatial variations in urban public ridership derived from GPS trajectories and smart card data. *Journal of Transport Geography* **2018,** 69, 45-57.
21. Zhou, M.; Wang, D.; Li, Q.; Yue, Y.; Tu, W.; Cao, R., Impacts of weather on public transport ridership: Results from mining data from different sources. *Transportation research part C: emerging technologies* **2017,** 75, 17-29.
22. Jendryke, M.; Balz, T.; McClure, S. C.; Liao, M., Putting people in the picture: Combining big location-based social media data and remote sensing imagery for enhanced contextual urban information in Shanghai. *Computers, Environment and Urban Systems* **2017,** 62, 99-112.
23. Kim, K.-S.; Kojima, I.; Ogawa, H., Discovery of local topics by using latent spatio-temporal relationships in geo-social media. *International Journal of Geographical Information Science* **2016,** 30, (9), 1899-1922.
24. Martí, P.; Serrano-Estrada, L.; Nolasco-Cirugeda, A., Social media data: Challenges, opportunities and limitations in urban studies. *Computers, Environment and Urban Systems* **2019,** 74, 161-174.
25. Kim, Y.-L., Seoul's Wi-Fi hotspots: Wi-Fi access points as an indicator of urban vitality. *Computers, Environment and Urban Systems* **2018,** 72, 13-24.
26. Shih-Lung, S.; Tsou, M.-H.; Xinyue, Y., Editorial: Human dynamics in the mobile and big data era. *International Journal of Geographical Information Science* **2016,** 30, (9), 1.
27. Yuan, M., Human dynamics in space and time: A brief history and a view forward. *Transactions in GIS* **2018,** 22, (4), 900-912.
28. Liu, Y.; Liu, X.; Gao, S.; Gong, L.; Kang, C.; Zhi, Y.; Chi, G.; Shi, L., Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers* **2015,** 105, (3), 512-530.
29. De Nadai, M.; Staiano, J.; Larcher, R.; Sebe, N.; Quercia, D.; Lepri, B. In *The death and life of great Italian cities: a mobile phone data perspective*, Proceedings of the 25th international conference on world wide web, 2016; 2016; pp 413-423.
30. Huang, B.; Zhou, Y.; Li, Z.; Song, Y.; Cai, J.; Tu, W., Evaluating and characterizing urban vibrancy using spatial big data: Shanghai as a case study. *Environment and Planning B: Urban Analytics and City Science* **2019**, 2399808319828730.
31. Long, Y.; Huang, C., Does block size matter? The impact of urban design on economic vitality for Chinese cities. *Environment and Planning B: Urban Analytics and City Science* **2019,** 46, (3), 406-422.
32. Wu, F., Emerging Chinese cities: Implications for global urban studies. *The Professional Geographer* **2016,** 68, (2), 338-348.
33. Zhao, Z.; Shaw, S.-L.; Yin, L.; Fang, Z.; Yang, X.; Zhang, F.; Wu, S., The effect of temporal sampling intervals on typical human mobility indicators obtained from mobile phone location data. *International Journal of Geographical Information Science* **2019,** 33, (7), 1471-1495.
34. Li, J.; Li, J.; Yuan, Y.; Li, G., Spatiotemporal distribution characteristics and mechanism analysis of urban population density: A case of Xi'an, Shaanxi, China. *Cities* **2019,** 86, 62-70.
35. Gehl, J., *Life between buildings: using public space*. Washington: Island press: 1971.
36. Jacobs, J., *The death and life of great American cities*. Vintage: 1961.

37. Glaeser, E., *Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Hap pier*. Penguin Press: 2012.
38. Storper, M., *Keys to the city: How economics, institutions, social interaction, and politics shape development*. Princeton University Press: 2013.
39. Wirth, L., Urbanism as a Way of Life. *American journal of sociology* **1938,** 44, (1), 1-24.
40. Simmel, G., The metropolis and mental life. In *The urban sociology reader*, Routledge: 2012; pp 37-45.
41. Belzer, D.; Autler, G., *Transit oriented development: moving from rhetoric to reality*. Brookings Institution Center on Urban and Metropolitan Policy Washington, DC: 2002.
42. Xia, C.; Yeh, A. G.-O.; Zhang, A., Analyzing spatial relationships between urban land use intensity and urban vitality at street block level: A case study of five Chinese megacities. *Landscape and Urban Planning* **2020,** 193, 103669.
43. Li, J.; Li, M.; Long, Y.; Dang, A., China polycentric cities based on baidu heatmap. *Shanghai Urban Planning Review* **2016,** 3, 30-36.
44. Tan, X.; Huang, D.; Zhao, X.; Yu, Y.; Leng, B.; Feng, L., Jobs housing balance based on Baidu thermodynamic diagram. *J. Beijing Normal Univ. Natural Sci.* **2016,** 52, (5), 622-627.
45. Wu, Z.; Ye, Z., Research on urban spatial structure based on Baidu heat map: A case study on the central city of Shanghai. *City Planning Review* **2016,** 40, (4), 33-40.
46. Lucang, W., Spatial-temporal characteristics of urban population aggregation based on baidu heat map in central areas of Wuhan city. *J. Hum. Settl. West China* **2018,** 33, 52-56.
47. Cattell, V.; Dines, N.; Gesler, W.; Curtis, S., Mingling, observing, and lingering: Everyday public spaces and their implications for well-being and social relations. *Health & place* **2008,** 14, (3), 544-561.

Chapter 7. Conclusion

7. 1 Conclusion

There is a complex inter-causal link between urban development and population mobility. The development of cities requires a steady stream of labor to support, which promotes large-scale population mobility. Sufficient labor has ensured the rapid development of cities, and the uneven regional development has become a new "pull force" for population mobility. Therefore, it is necessary to study the two-way connection between the urban development and population to provide theoretical support for China's urbanization process in the future.

This paper combines official statistics data and emerging big data, using spatial data analysis methods to study the interactive relationship between urban development and population mobility at different scales. The main work and results can be summarized as follows:

In *chapter one*, BACKGROUND AND PURPOSE OF THE STUDY. Research background and significance of urban development and population mobility is demonstrated. In addition, the previous study about urban development and its relationship with population mobility this is reviewed. Then the purpose of the study is proposed.

In *chapter two*, DATA DESCRIPTION AND RESEARCH METHODOLOGY. First, we introduced the theory and application progress of Geographic Information System (GIS). Second, we introduced GIS-based spatial data analysis methods and spatial econometric models that capture spatial effects. Then, combining different research contents, we also explained in detail the theory and applications of social network analysis and logistic regression models. Finally, we focused on the data used in the paper, including official statistics data, Tencent location big data, Baidu heat map data, and national floating population dynamics monitoring survey data.

In *chapter three*, SPATIOTEMPORAL CHARACTERISTICS AND DRIVING FORCES OF POPULATION MOBILITY IN URBAN CHINA FROM 2000-2018. Based on official statistics data, we studied the characteristics of population mobility in China's 290 prefecture-level cities and their driving forces. Specifically, during 2000-2010, the destinations of population migration were concentrated in the Pearl River Delta, the Yangtze River Delta, and the Beijing-Tianjin region, while the inland cities in the central and western regions were lost due to large numbers of population outflows. However, after 2010, The scale of population migration in the eastern coastal mega cities is gradually slowing down, while the phenomenon of population emigration even appears in the surrounding second and third- tier cities. Meanwhile, the central and western inland capital cities and regional central cities have large-scale population immigration, the population immigration in the second and third-tier hinterland cities is also obvious. At the same time, we also analyzed the impact of urban inhabitant environment, social economy and public policies on population mobility. Although economic differences are still the main factor in population mobility, it seems that the impact of inhabitant environment is becoming more and more important. In addition, public policies have a weak role in promoting population mobility in megacities. This part of the research systematically analyzes and summarizes the population mobility in China in the past, and explores its impact on population mobility from three different aspects of urban development.

In *chapter four*, SPATIOTEMPORAL PATTERNS OF POPULATION MOBILITY AND ITS DETERMINANTS IN URBAN CHINA DURING SPRING FESTIVAL OF 2019. With the help of Tencent location big data, we can capture the time and space trajectory of population mobility, which provides us with necessary data support for studying the spatial and temporal patterns of population mobility. In this part, we use high-precision space-time big data to describe the network pattern of population mobility during the Spring Festival (China's most important festival). We found that the population mobility network during the Spring Festival presents a diamond structure. The four vertices of the diamond are the core cities of the four major urban agglomerations. The population mobility network presents "small world" characteristics. With the help of a semiparametric geographically weighted regression model, we focused on analyzing the impact of the urban economy on population mobility. Wage levels, the development of the tertiary industry, and investment in foreign capitals are important factors in attracting population mobility between regions. Our results provide a good supplement to the inter-regional population research. After considering the spatial heterogeneity, it provides reference for urban development in different regions and cities.

In *chapter five*, EMPIRICAL ANALYSIS ON THE MECHANISM OF POPULATION MOBILITY PROMOTING URBAN DEVELOPMENT IN CHINA. In this chapter, we empirically analyze the impact of population mobility on regional or urban development and its transmission mechanism at the regional scale. We selected the areas with the largest population mobility scale and the best economic development, and the Yangtze River Delta urban agglomeration for empirical analysis. Our results show that within the Yangtze River Delta urban agglomeration, population mobility promotes regional and urban development, which is manifested in the rapid growth of urban economy. The panel data model also emphasizes that large-scale population influx can transform the industrial structure, improve the level of technological innovation, increase household consumption, and improve the labor employment market, thereby boosting the development of regions or cities.

In *chapter six*, SPATIOTEMPORAL DYNAMICS OF POPULATION MOBILITY AND MECHANISM ANALYSIS AT STREET BLOCK LEVEL: A CASE STUDY IN QINGDAO, CHINA. In this chapter, we use microscopic urban streets as the research object to describe the temporal and spatial dynamics of population mobility within the city and the impact of the built environment. Baidu heatmap data provides us with real-time urban population distribution. The temporal and spatial differences of urban population distribution reflect the dynamic characteristics of urban vitality, and in turn reflect the changing trend of population mobility within the city. Our results show that the population mobility in the city has a certain regularity. During the weekdays, it is shown as shrinking to the regional center or city center at daytime, and spreading to the surrounding areas at nighttime, and vice versa on weekends. The regression model shows that there is a significant spatial-temporal correlation between the urban built environment (density, diversity, design, distance to transit and destination accessibility) and the population mobility. Our results provide a reference for urban design in maintaining urban vitality and attracting population mobility, which is crucial for the sustainable development of regions and cities.

In *chapter eight*, CONCLUSION. The whole summary of each chapter has been presented.