

Analysis of the Characteristics of City Scale Distribution and Evolutionary Trends in China

Min Zhang*, Zhen Jia

College of Science, Guilin University of Technology, Guilin, China

Email: *2351007655@qq.com, jjzzz0@163.com

How to cite this paper: Zhang, M. and Jia, Z. (2021) Analysis of the Characteristics of City Scale Distribution and Evolutionary Trends in China. *Open Journal of Statistics*, 11, 443-462.

<https://doi.org/10.4236/ojs.2021.113028>

Received: June 2, 2021

Accepted: June 22, 2021

Published: June 25, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Based on the urban resident population statistics from 2005 to 2018, this paper analyzes the distribution and evolution of city scale in China by screening city samples according to the threshold criteria and using empirical research methods such as the City Primacy Index, the Rank-Scale Rule, the Gini coefficient of city scale, Kernel Density Estimation and Markov transfer matrix. The results show that: the most populous city in China has obvious advantages. The population distribution is concentrated in high order cities and in accordance with the law of order-scale; the economic scale of cities is in a concentrated state, the gap between the economic development levels of different types of cities is large, and the megacities are more attractive, which to a certain extent limit the development of the scale of the rest of the cities; the number of China's city population is increasing, however, the gap between the population scale of other cities and the most populous city continues to be large, and the structure of city population scale is not reasonable enough; megacities and megalopolises keep their original scale levels unchanged to a large extent, and the scale transition between the two types of cities is rather difficult. Finally, based on the explanatory framework of the dynamics of city scale evolution, policy recommendations are proposed to promote a more balanced distribution of city scale.

Keywords

City Scale, The City Primacy Index, The Rank-Scale Rule, The Gini Coefficient of City Scale, Kernel Density Estimation

1. Introduction

City scale is a core topic in city research, and an accurate understanding of the evolutionary dynamics of China's city scale can provide strong technical support

for sustainable city development and high efficiency in the use of public resources [1]. At the current stage, the uncontrolled population growth and the land resources not available to match in some cities in China have seriously threatened the healthy development of cities to a certain extent [2]. Therefore, exploring the distribution characteristics of China's city scale, studying their evolutionary trends, and reasonably predicting the expansion trend and spatial distribution characteristics of city population scale will help promote the optimization and upgrading of China's city spatial structure, as well as enhance the capacity of city services and promote the shift of China's regional economy to high-quality development.

There are countless studies on the characteristics of city scale distribution at home and abroad. Jefferson, a foreign scholar, first introduced the city primacy in 1931 when he discussed the characteristics of regional city distribution [3]; Singer relates city rank order to city scale distribution, revealing a quantitative relationship between them [4]; Marshall used the city Gini coefficient calculation method to study the concentration of city scale distribution [5]. In China, Xinghua Feng *et al.*, used Kernel density analysis to study the evolution of city scale hierarchy in the Yangtze River Economic Zone [6]; Wei Chunjiang *et al.*, applied the city place-order-scale rule to study traditional farming villages with a new perspective [7]; Wan Qing *et al.*, argued that factors such as geographic location, level of economic development and political system are responsible for the dramatic expansion of the scale of large cities [8]; Li, S. *et al.* explored the relationship between city scale and the level of economic development using Zipf's law, arguing that the factors affecting economic development also have an impact on the evolution of city scale distribution trends [9]; Sen-Sheng Li *et al.*, applied a joint cubic equation model in economics and explored that the city-rural income gap also has an impact on the city scale distribution [10]; Yan Zhou *et al.*, improved the land use model for city scale development by adding land constraint parameters in terms of population growth and land expansion to derive a quantitative relationship between the rate of land expansion and the rate of population growth [11]. Meanwhile, some scholars have explored the characteristics of city scale distribution in different regions in China, such as the Poyang Lake City Cluster [12] and the Central Liaoning City Cluster [13], and concluded that the city scale of each region in China shows uneven distribution.

However, there are fewer relevant studies using multiple research methods such as Zipf's law [14], kernel density estimation [15], and Markov transfer matrix [16] [17] in the aforementioned studies, which, together with the lack of accuracy of indicators characterizing city scale in empirical studies on the distribution and dynamic evolution of city scale in China, and the limitations of administrative statistical units, lead to some discrepancies between the estimated and true city scale. Therefore, this paper will consider the distribution characteristics of China's city scale from multiple dimensions using multiple indicators and explore the dynamic evolution trend of China's city scale vertically from the time

series dimension. This paper selects Chinese city data from 2005-2018 and divides them into four-time points with five-year plans to analyze the distribution and dynamic evolution patterns of city scale.

2. City Scale Research Methods

2.1. City Primacy

The American geographer Jefferson (1939) considered the first city as the most populous city in a region and its number would be higher than the second city by two times, therefore, the ratio of the population scale of the first and second cities was proposed to represent the firstness, *i.e.*, S_2 . Later on, some scholars improved it and proposed the four-city index (S_4) and the eleven-city index (S_{11}) [18], which are calculated as follows:

$$\begin{cases} S_2 = \frac{P_1}{P_2} \\ S_4 = \frac{P_1}{P_2 + P_3 + P_4} \\ S_{11} = \frac{2P_1}{P_2 + P_3 + \dots + P_{11}} \end{cases} \quad (1)$$

where P_1, P_2, P_3, \dots denotes the number of population of cities in order 1st, 2nd, 3rd, ..., respectively. According to Felix Auerbach's (1913) Theory of City Rank Scale Distribution [19], ideally S_2 is 2, and S_4, S_{11} are 1.

2.2. The Rank-Scale Rule test

The American linguist George Kingsley Zipf (1935) established the city Rank-Scale Rule, known as the Zipf distribution. The general relationship equation is as follows:

$$P_i = P_1 \times R_i^{-q} \quad (2)$$

To transform the expression (2), we get:

$$\ln P_i = \ln P_1 - q \ln R_i \quad (3)$$

where: P_1 is the population of the most populous city; P_2, P_3, P_4, \dots refers to the population of the 2nd, 3rd, and 4th, etc., cities, respectively; R_i represents the order of i city; q denotes the Zipf index.

2.3. City Gini Coefficient

Marshall first introduced the city Gini coefficient for the study of city scale, and the specific formula is:

$$G = T/2S(n-1) \quad (4)$$

Specifically, n represents the number of cities. G represents the city Gini coefficient, which can be expressed as PG or EG. PG refers to the Gini coefficient of city population scale, with $S = \sum_{i=1}^n P_i$ and $T = \sum_{i,j=1, i \neq j}^n |P_i - P_j|$ in expression (4)

respectively representing the total city population in the region and the sum of the absolute values of the difference between the population scale of any two cities. EG denotes the Gini coefficient of city economic scale. $S = \sum_i^n GDP_i$ and $T = \sum_{i,j=1, i \neq j}^n |GDP_i - GDP_j|$ in expression (4) are the total city economic scale in the region and the sum of the absolute values of the difference between the economic scale of any two cities, respectively.

2.4. Kernel Density Estimation (KDE)

Rosenblatt (1955) and Emanuel Parzen proposed the Kernel Density Estimation, also known as KDE, in 1962, which is a nonparametric estimation. The general expression is:

$$f(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (5)$$

where $f(x)$ is the probability density function to be estimated and $k(\cdot)$ is the kernel function, satisfying symmetry and $\int K(x)dx = 1$, also, n is the number of observations and h is the bandwidth. It is worth noting that the choice of h affects the smoothness of the estimated function as well as the fitness of the model, which in principle should satisfy the minimum of the mean square error. In this paper, the commonly used Epanechnikov kernel function is selected, and the h -value is automatically adjusted in MATLAB according to the sample data characteristics to be optimal.

2.5. Markov Transfer Matrix

The city scale data are divided into k classes, and to calculate the probability of cities in each class and the probability of transferring cities between each class, this process can be approximated as a Markov process.

The Markov transfer matrix is denoted by M , M_{ij} denotes the one-step transfer probability of a city belonging to rank i in time t converting to rank j in the next period $t+1$, i.e., $M_{ij} = n_{ij}/n_i$, where n_{ij} denotes the sum of the number of cities converting from rank i in time t to rank j in time $t+1$. n_i is the sum of the number of all rank i cities in time t . The process of city scale evolution is accompanied by the rise and fall of each city scale class, i.e., the change of mobility in different classes of cities. In this paper, two mobility measures are used, the first of which is the SMI (Shorrocks Mobility Index) as defined by Shorrocks [20]:

$$SMI = \frac{k - \sum_{i=1}^k M_{ii}}{k-1} \quad (6)$$

The scale of SMI depends on $\sum_{i=1}^k M_{ii}$. The smaller the SMI value, the weaker the mobility, and the greater the probability that a city belonging to a certain class in time t will still belong to the same class in time $t+1$; on the contrary,

the larger the calculated SMI value, the stronger the mobility.

The second one is to construct a new liquidity measure DMI (Direction Mobility Index). It can be expressed as:

$$DMI = \sum_{i=1}^k 2^{i-j-1} \left(\sum_{j=i+1}^k M_{ij} - \sum_{j=1}^{i-1} M_{ij} \right) \quad (7)$$

The former part of expression (7) indicates the mobility index of cities moving to higher levels, *i.e.* upward mobility level, and the latter part indicates downward mobility level, where 2^{i-j-1} is the weight of mobility across levels, when the higher the number of levels crossed, the more hard it is to move, so it needs to be given a greater weight. $DMI > 0$ indicates that cities are more likely to cross to higher levels. $DMI < 0$ indicates that cities are more mobile downward, *i.e.*, cities are underdeveloped in later stages and their scale levels are decreasing.

3. Sample Selection and Data Sources

Most studies on city scale use two types of data, one is the data of “population of municipal districts” in the City Statistical Yearbook, which includes the entire population of city areas and a large number of rural areas and is easy to overestimate city scale; the other type of data is the statistical caliber of household population, using the data of “city non-agricultural population” to measure, which lacks systematic and accurate statistics of the transient and mobile population, it is easy to underestimate city scale. Accordingly, the choice of either “municipal population” or “non-agricultural population” can lead to errors in measuring the scale of cities. The city resident population of the national census data from 2005 to 2018 is selected in this paper, and in order to more accurately portray the mobility characteristics of cities, a population scale of 5 million and the National GDP per capita (billion yuan) of that year are also selected as the threshold lower bound at the same time. Defining thresholds is of high significance for the study: for one, this paper uses both the population of 5 million and the National GDP per capita as the lower threshold to filter out cities that meet both the population scale and a certain level of economic development, so that it is easy to identify them from other cities and the economic links are more obvious. The second one is that cities will gradually prosper or decay in the process of development, and setting the threshold can make the cities whose original resident population scale exceeds the threshold enter the sample, while those below the threshold in the later stage drop out of the sample.

In this paper, the population data from 2005 to 2018 are obtained from the China City Statistical Yearbook 2006-2019, China Statistical Yearbook 2006-2019 and China Population and Employment Statistical Yearbook 2006-2019. The three major indicators of resident population and gross product, and gross product per capita are selected as the research data by combining the above yearbooks, and the characteristics and evolutionary trends of city scale distribu-

tion in China are studied by using MATLAB, ArcGIS and other measurement and statistical analysis software.

4. Analysis of City Scale Distribution Characteristics

4.1. City Scale Distribution

In this paper, China's city resident population data from 2005 to 2018 are screened for sample cities based on threshold conditions, and the current city scale class is also used as the basis for classifying the city scale class. At present, the State Council stipulates the criteria for dividing city scale as: 1) megacity (population number over 10 million); 2) megalopolis (population number 5 to 10 million); 3) large city (population number 1 to 5 million, among which the population number of type I large cities is 3 to 5 million, and the population number of type II large cities is 1 to 3 million); 4) medium city (population number 500,000 to 1 million); 5) small city (population number less than 500,000, including type I small cities with population number 200,000 to 500,000, and type II small cities with population number less than 200,000), according to which the distribution of city population class scale is obtained in this paper (See [Table 1](#) and [Figure 1](#)).

As seen in [Table 1](#), there are more megalopolis than megacities in the sample period. Comparing the scale of cities in different periods shows that the number of megacities has had a rapid growth in recent years in the period from 2010 to 2015, with little change in the number of megalopolis, and the mobility between cities is well reflected due to the establishment of threshold conditions. Chongqing's resident population exceeded 30 million in 2015, making it the megacity with the

Table 1. Distribution of city class scale.

Scale	2005	2010	2015	2018
Megacities (Population > 10 million)	4 Shanghai, Beijing, Chengdu, Tianjin	6 Shanghai, Beijing, Tianjin, Chengdu, Guangzhou, Shijiazhuang	10 Chongqing, Shanghai, Beijing, Tianjin, Chengdu, Guangzhou, Shenzhen, Shijiazhuang, Suzhou, Wuhan	10 Chongqing, Shanghai, Beijing, Chengdu, Tianjin, Guangzhou, Shenzhen, Wuhan, Suzhou, Zhengzhou
	33	35	33	33
Megalopolises (5 million < Population < 10 million)	Harbin, Shijiazhuang, Wuhan, Weifang, Shenzhen, Qingdao, Jining, Xi'an, Quanzhou, Wenzhou, Guangzhou, Suzhou, Nantong, Changchun, Tangshan, Shenyang, Cangzhou, Fuzhou, Hangzhou, Dongguan, Zhengzhou, Yantai, Luoyang, Changsha, Kunming, Jinan, Nanjing, Foshan, Dalian, Taizhou, Ningbo, Dezhou, Tai'an	Suzhou, Harbin, Wuhan, Weifang, Shenzhen, Xuzhou, Qingdao, Hangzhou, Xi'an, Zhengzhou, Jining, Shenyang, Nanjing, Quanzhou, Wenzhou, Changchun, Tangshan, Ningbo, Yancheng, Nantong, Cangzhou, Yantai, Fuzhou, Changsha, Jinan, Dalian, Luoyang, Dongguan, Kunming, Wuxi, Foshan, Taizhou, Tai'an, Jinhua, Nanchang	Harbin, Zhengzhou, Weifang, Wenzhou, Qingdao, Hangzhou, Xi'an, Xuzhou, Quanzhou, Dongguan, Shenyang, Nanjing, Ningbo, Tangshan, Hefei, Changchun, Fuzhou, Foshan, Changsha, Nantong, Yancheng, Jinan, Yantai, Dalian, Luoyang, Kunming, Wuxi, Taizhou, Yueyang, Xiangyang, Tai'an, Jinhua, Nanchang	Xi'an, Hangzhou, Harbin, Weifang, Qingdao, Wenzhou, Xuzhou, Quanzhou, Nanjing, Dongguan, Shenyang, Ningbo, Changsha, Hefei, Tangshan, Foshan, Fuzhou, Changchun, Jinan, Nantong, Yancheng, Yantai, Dalian, Luoyang, Kunming, Wuxi, Taizhou, Xiangyang, Tai'an, Jinhua, Nanchang, Zhangzhou, Shaoxing

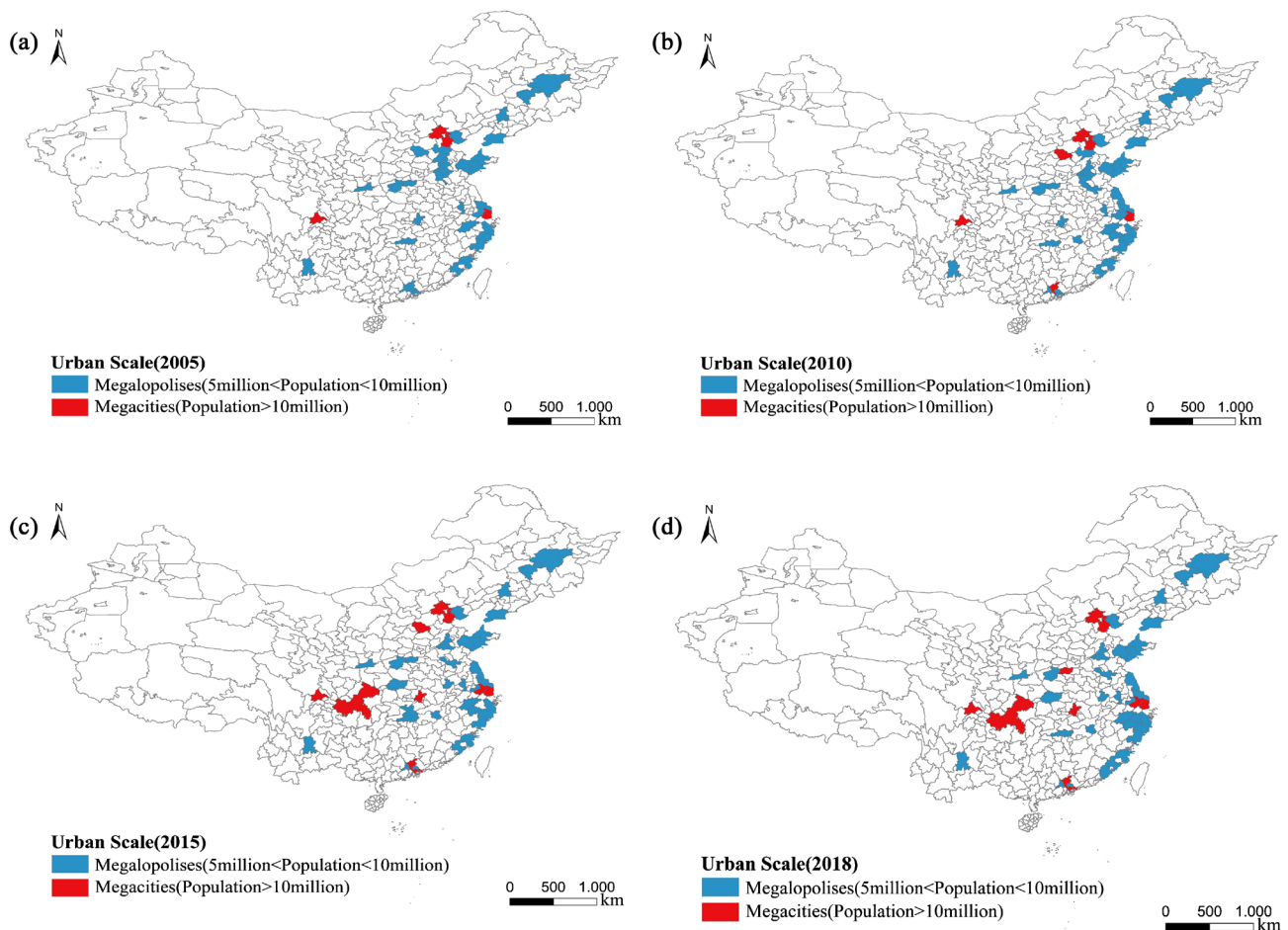


Figure 1. Visualization of City Class Scale Distribution. (a) City Scale Distribution in 2005; (b) City Scale Distribution in 2010; (c) City Scale Distribution in 2015; (d) City Scale Distribution in 2018.

largest resident population in China. Between 2015 and 2018, Shijiazhuang withdrew from the sample because its per capita gross product was lower than the national per capita gross product in the same year, which was due to the unhealthy industrial structure, relying mainly on light industry, and most people working in purely manual technical categories, which did not have significant advantages compared to other industries. Moreover, the foundation of the Internet industry is quite weak. The most critical thing is that Shijiazhuang has only one main line of Beijing–Guangzhou Railway in its territory, and as the capital city of the province, it has no advantages in terms of geographical location, further leading to Shijiazhuang’s lack of development momentum out of the sample.

Following the visualization of the data in [Table 1](#), [Figure 1](#) is available, from which we can observe the distribution areas of cities with large city population and economic scale in China. From [Figure 1](#), we can see that megalopolises are widely distributed, mainly in the eastern and southeastern coastal ports, followed by the Central Plains and Northeast China, and the Southwest China with only two provincial capitals and a municipality, while none of the cities in the West

and Northwest China were selected. Geographically, the new megacities added between 2005 and 2018 are distributed in South, Southwest and Central China, among which the number of megacities belonging to Southwest and Central China is high, and the megacities show a trend of development from the eastern coastal region to the inland.

In order to directly observe the distribution areas of cities with a population of 10 million or more and a citywide GDP in the top ten of the national GDP during the sample period, we imported the city population scale and GDP data into ArcGIS 10.2 software separately for visual analysis (Figure 2).

Comparing Figure 2(a) with Figure 2(b), it can be seen that the distribution of city population scale in China is basically consistent with the distribution of GDP level. Most of the cities in the top 10 in China in terms of population are distributed in the eastern, the southeast coast, and the southwest and central plains, while there are none in the western areas of China. This is due to the reform and the policy of opening up to the outside world have started from the eastern coastal areas, first get the advantage of development and have a stronger multi-functional development potential, as well as economic and technological development started earlier, and the implementation of different preferential policies with the inland, moreover, foreign trade, sea transport make the eastern coast easier to attract foreign investment, so that the eastern commodity economy developed. However, most of the western area of China is a plateau with complex topography, coupled with poor transportation and backward city public services, the gap between the industrial gradient in the eastern region is large, which seriously restricts the migration of population.

According to the population scale distribution from 2005 to 2018 in China, the percentage of the number of cities in each class is plotted (Figure 3). After 13 years of economic development, the number of megacities and megalopolis evolved from 89.2% and 10.8% to 76.7% and 23.3% respectively. As seen in Figure 3, the number of megacities is gradually increasing and the number of megalopolis

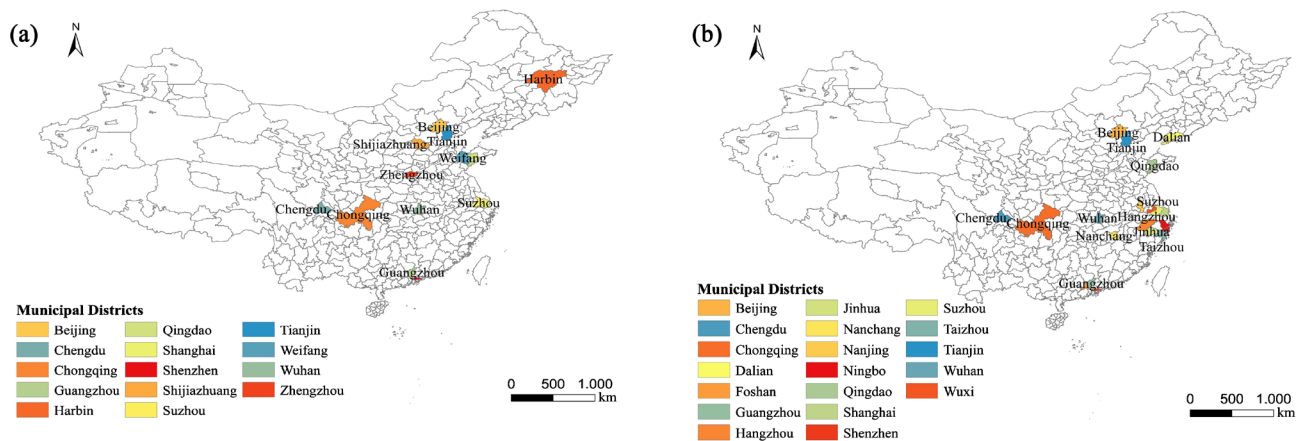


Figure 2. Distribution of Cities with City Population of 10 Million or More and the Top 10 Cities in the Country in terms of City GDP, 2005 to 2018. (a) Distribution of Cities with Population of 10 Million or More; (b) The Top Ten Cities in the National GDP Distribution Area.

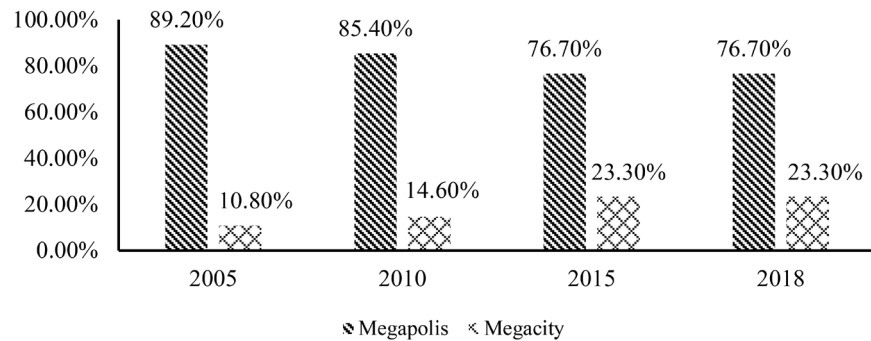


Figure 3. Percentage of the Number of Megacities and Megalopolis in the Country from 2005 to 2018.

is gradually decreasing. Combining **Figure 1** and **Figure 2**, we can see that the population distribution is spatially consistent with the level of economic development, and the cities with high economic level also have more population, *i.e.*, the mobile population gradually flocks to the cities with developed economy and convenient transportation.

4.2. Evolution of the City Primacy Index

Table 2 reflects the order and population scale of some cities in China, and it can be seen that from 2005 to 2010, the population scale of Guangzhou grew the fastest, followed by Suzhou. Guangzhou has become an important port for foreign trade and material exchange in China after the reform and opening. Moreover, it is a key and pioneer city for investment and business attraction, with a developed economic level that attracts a large number of people. Suzhou is the city that has been most affected by the economic diffusion of Shanghai, and its economy has grown so rapidly that it has become the second largest industrial city in China after Shanghai in just a few years, making it a city with a rapid population growth. From 2010 to 2015, the implementation of the “industrial city” strategy to further promote the optimization and upgrading of Chongqing’s industry, so that it had formed a modern industrial system supported by multi-pillar industries. In addition, the entry of a considerable number of multinational corporations and state-owned enterprises, as well as the large number of international industrial transfers that Chongqing has received during this period, have attracted a large inflow of mid- to-high-end talents, and the overall level of economic development in the city has increased rapidly as well as the per capita GDP, further contributing to the city’s continued expansion. After 2015, Chongqing completed the diversification of its economic industry chain, while the continuous improvement of the city’s infrastructure and public services has led to increasing economic activity, and in 2018 Chongqing remained the city with the largest resident population in the country. Meanwhile, Chengdu, which is located in the same southwest region as Chongqing, has also introduced a series of policies to attract outstanding talents and business resources to enter. Besides, in early 2014, some data indexes such as number of luxury brands, scale of

Table 2. Order and Population Scale of Some Cities in China.

Order	2005		2010			2015			2018		
	city	population	city	population	increase	city	population	increase	city	population	increase
1	Shanghai	1778	Shanghai	2256	0.27	Chongqing	3004	0.05	Chongqing	3088	0.03
2	Beijing	1515	Beijing	1858	0.23	Shanghai	2420	0.07	Shanghai	2421	0
3	Chengdu	1207	Tianjin	1263	0.22	Beijing	2161	0.16	Beijing	2162	0
4	Tianjin	1033	Chengdu	1144	-0.05	Tianjin	1531	0.21	Chengdu	1618	0.11
5	Harbin	970	Guangzhou	1037	0.39	Chengdu	1454	0.27	Tianjin	1558	0.02
6	Shijiazhuang	956	Shijiazhuang	1003	0.05	Guangzhou	1329	0.28	Guangzhou	1470	0.11
7	Wuhan	852	Suzhou	992	0.34	Shenzhen	1107	0.24	Shenzhen	1277	0.15
8	Weifang	851	Harbin	992	0.02	Shijiazhuang	1065	0.06	Wuhan	1098	0.05
9	Shenzhen	814	Wuhan	944	0.11	Suzhou	1060	0.12	Suzhou	1070	0.01
10	Qingdao	812	Weifang	902	0.06	Wuhan	1047	0.11	Zhengzhou	1000	0.06
11	Jining	804	Shenzhen	896	0.1	Harbin	974	-0.01	Xi'an	981	0.13

stores and consumption amount in Chengdu showed excellent development, thus Chengdu was named as “the third city of luxury in China”, ranking after Beijing and Shanghai, which shows that the population and economic scale of Chengdu are rapidly developing. When combined with [Table 1](#) and [Table 2](#), it can be concluded that Zhengzhou has developed steadily and continuously ascending from a medium-scaled megalopolis to a megacity. Zhengzhou has been the transportation hub of the south and north from the ancient times, and has unparalleled geographical advantages. The implementation of a diversified economic layout in recent years has led to the development of heavy industry, primary manufacturing, and labor-intensive enterprises in Zhengzhou, attracting a large number of mid- to-high-end talents, making it the fastest-growing GDP and fastest-improving economic ranking of any major northern city. Throughout the sample period, Guangzhou, Shenzhen has developed the most momentum and maintained a high growth rate of continuous growth, which is inextricably linked to their industrial structure that emphasizes science and technology innovation.

Equation (1) was further applied to calculate the trends of S_2 , S_4 and S_{11} from 2005 to 2018 population data (see [Figure 4](#)). As can be noted from [Figure 4](#): S_2 is significantly higher than the theoretical state 2, S_4 is close to the theoretical value 1, and S_{11} is significantly lower than the theoretical value 1 from 2005 to 2018. This happens as a result of the denominator value of the summed eleven-city index being too large, making the eleven-city index value much smaller than ideal. It means that the population scale stratification among cities is unreasonable, and the population distribution is concentrated in high order cities, and the most populous city has obvious advantages of strong attraction. Shanghai, Beijing and other high order cities have obvious population clustering effect,

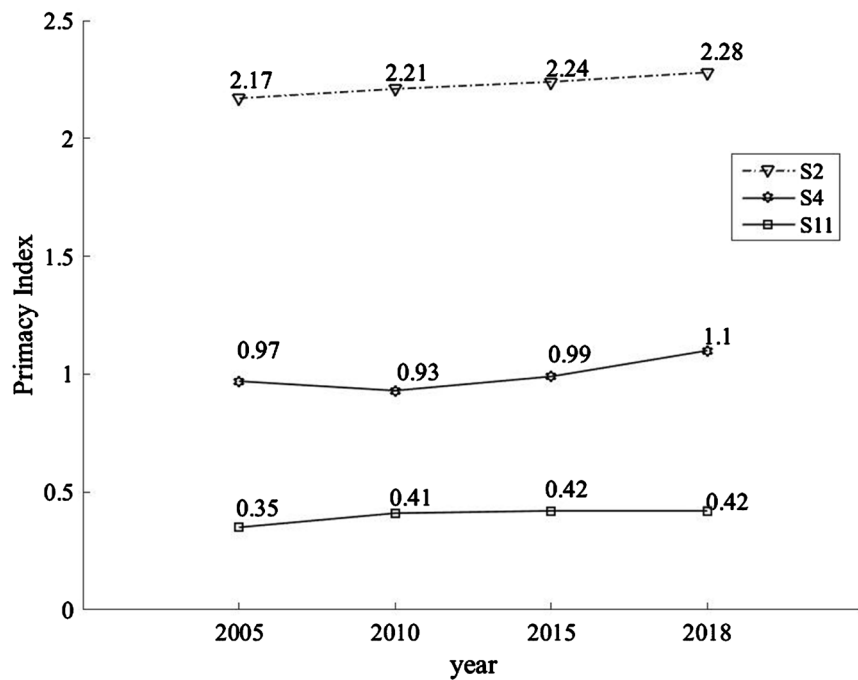


Figure 4. Changes in City System Primacy Index from 2005 to 2018.

and the “double core” spatial pattern is remarkable, showing the unbalanced development trend of city population scale system. In recent years, Chengdu, Shenzhen, Guangzhou, Xi’an and other high order cities have experienced rapid economic scale development, which have increased their attractiveness to other cities and further contributed to the continued expansion of population scale, coupled with the current implementation of a series of talent policies in these cities, the future eleven-city index will either stabilize or decline. At the same time, looking at the trend of changes throughout the sample period, S_2 and S_4 are increasing year by year, and the reason for this phenomenon is that Shanghai, the most populous city from 2005 to 2015, and Chongqing, the most populous city from 2015 to 2018, continue to expand their city scale, causing a gradual strengthening of the return effect, which has a dampening effect on the scale development of high order cities such as Beijing, Tianjin, Chengdu, and Guangzhou.

Generally speaking, Shanghai, Chongqing, Beijing, Tianjin, Chengdu, Guangzhou and other high order cities have fast economic development and strong vitality, however, it will take a long time for the national population growth to achieve a balanced evolution due to the relatively large gap in city population scale development between cities in the Midwest and Northeast and other cities.

4.3. City Rank-Scale Rule test

Bringing the population data into Equation (3): $\ln P_i = \ln P_1 - q \ln R_i$, the Zipf distribution-related indicators were calculated to derive the trend of city scale distribution by year regression results from 2005 to 2018 (Figure 5).

From Figure 5, it can be noticed that residual sum of squares in regression

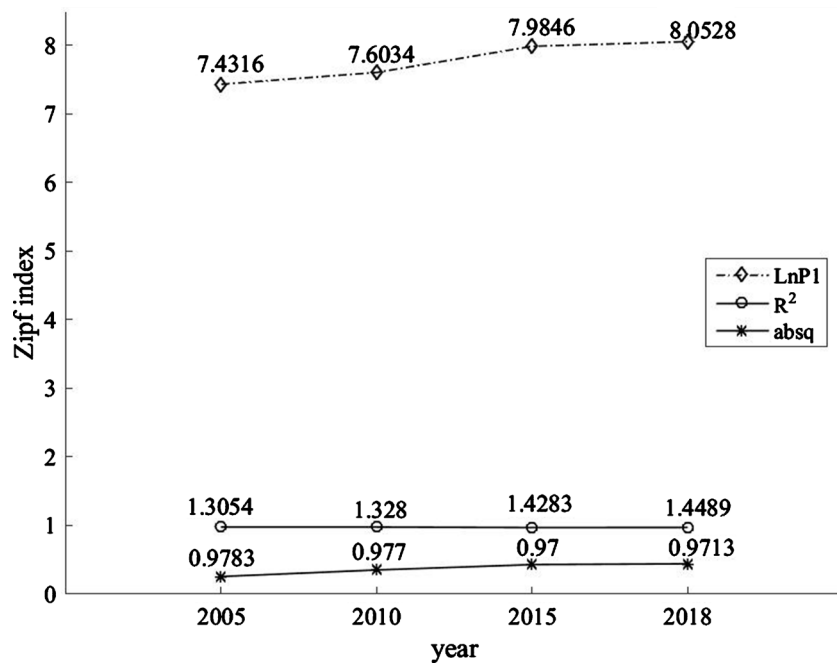


Figure 5. City System Rank-order-scale Rule Change from 2005 to 2018.

analysis is small from $R^2 > 0.9$, indicating that the regression model based on the rank-scale rule fits the actual city population data effectively. Meanwhile, by $|q| > 1$ can be indicated that the city scale distribution is in the concentration stage. Also represents that high order cities such as Shanghai, Beijing, Guangzhou, Shenzhen, etc., have strong economic, technological power and significant advantages, which is why the concentration of population scale is obvious, while low order cities such as Zhangzhou, Xiangyang, Yueyang, etc., have smaller scale and lower development level. As a whole, $|q|$ shows a rising trend, suggesting that the population scale of high order cities expands faster than that of low order cities, giving rise to a trend of large-scale dispersion and local concentration of city scale distribution. The strong attractiveness of some megacities has produced a monopoly effect in some ranges, exerting a driving effect while also producing a restrictive constraint effect, which results in a large difference in the speed of development between the high order cities and the rest, causing a disproportionate city development.

4.4. Analysis of City Scale Distribution Based on Gini Coefficient

Application of Equation (4) to calculate the Gini coefficient of population scale (PG) and the Gini coefficient of economic scale (EG) from 2005 to 2018 (Figure 6), the results show that:

First of all, according to the relevant United Nations organizations: if the Gini coefficient is less than 0.2, it refers to an absolutely balanced distribution; 0.2 to 0.3 represents a relatively balanced distribution; between 0.3 and 0.4 represents a relatively reasonable distribution; 0.4 to 0.5 represents a large gap in scale; and above 0.5 represents a significant gap in scale. The Gini coefficient of population

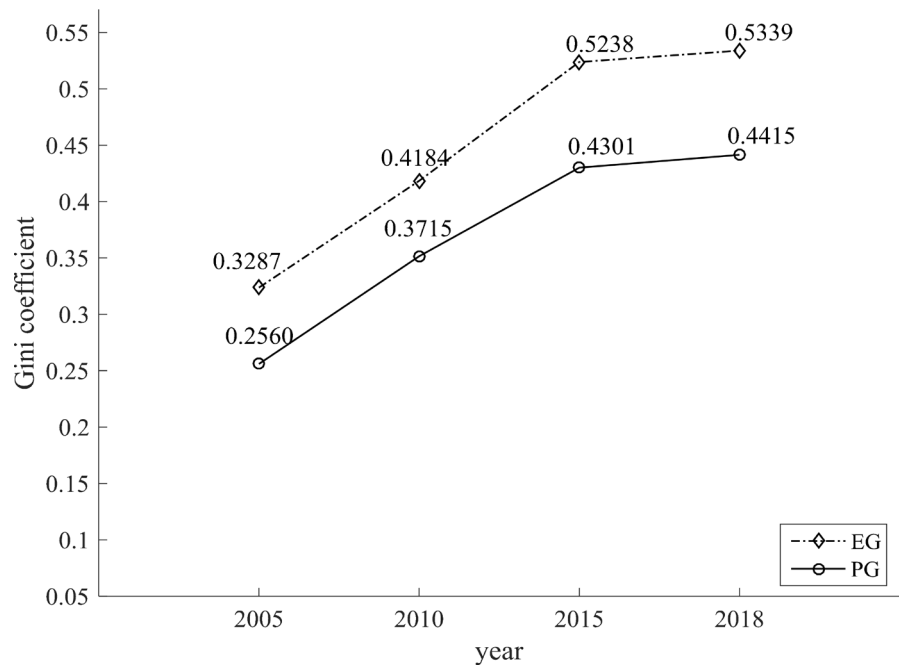


Figure 6. Change in Gini Coefficient of City Scale from 2005 to 2018.

scale from 2005 to 2010 is lower than 0.4, which implies a relatively reasonable distribution of population scale, however, the Gini coefficient of population scale from 2010 to 2018 is higher than 0.4, indicating a large gap between the population scale of megacities and megapolises. Overall, the distribution of city population scale shows a dispersion and a continuous tendency to regional concentration. From 2005 to 2015, the most populous city, Shanghai, whose diffusion effect has not been effectively utilized, did not have the resources of other cities effectively allocated, making the development of these cities slow and hindering the coordinated development among cities, making the Gini coefficient of population scale continuously larger. Due to the implementation of the national urbanization strategy of small and medium-scaled cities, small and medium-scaled cities have been developed rapidly, while the process of development of high-order cities is slow, thereby facing the challenge of coordinated development of city scale in China.

Secondly, the city GDP is brought into Equation (4) to calculate the city economic scale Gini coefficient, and it can be found from **Figure 6**, the city economic scale Gini coefficient is larger than the population scale Gini coefficient during the sample period, as well as the economic scale Gini coefficient gradually increases with the transition of time. This illustrates the significant economic agglomeration of megacities and the significant differences in the economic development of different cities, which are getting more and more obvious.

Eventually, a comprehensive comparison of the Gini coefficients of city population and economic scale in the same year reveals that the population and economic scale of each city are in a state of overall dispersion and small regional concentration.

4.5. Kernel Density Distribution of City Scale

We estimated kernel density for city scale data in 2005, 2010, 2015 and 2018 and obtained city scale kernel density distribution (Figure 7) by using MATLAB software with the horizontal coordinate being the city population scale and the vertical coordinate being the Kernel density, the results show that:

The first one is that the kernel density curve of city scale gradually shifts rightward with the increase of time, which represents the high urbanization rate and rapid growth of city population scale in China.

Moreover, when the peak of the kernel density curve is higher and the curve on the right side of the peak is steeper, it indicates that the distribution of the data is more concentrated [21]. From Figure 7, we can see that the peak of the kernel density curve of city scale shows a trend of gradual decrease, reflecting that the population scale of some megalopolises are increasing year by year, with a narrowing gap with the megacities over the years.

Last but not least, the kernel density curve in the figure is mainly decentralized single-peaked, and the peaks show a tendency to spread, which demonstrates that the population tends to be more and more regionally concentrated and the scale distribution among cities is uneven, however, the change in the number of megalopolises also reflects that this uneven trend is being gradually adjusted.

4.6. Dynamic Evolutionary Analysis of City Scale

In this paper, the growth characteristics of cities of different scale classes are considered on the basis of Markov matrix principles, and an in-depth dynamic analysis of the status of city scale distribution is carried out. Firstly, all cities in our selected sample are classified into two types: the first is type I for cities with a population scale of 5 - 10 million; the second is type II for cities with a population

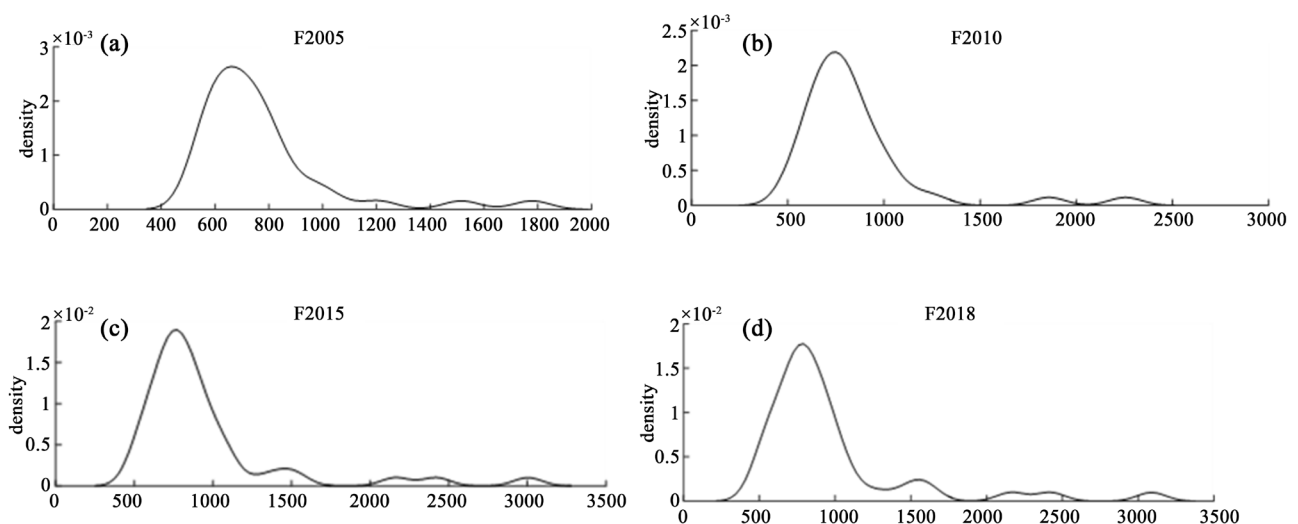


Figure 7. Change in Kernel Density Estimation Curve from 2005 to 2018. (a) Nuclear Density Curve in 2005; (b) Nuclear Density Curve in 2010; (c) Nuclear Density Curve in 2015; (d) Nuclear Density Curve in 2018.

scale of 10 million or more. After that, the probability of each city converting to a different class is calculated. Finally, the dynamics of city scale over the whole sample period is derived. Simultaneously consider the probability of entering and exiting the sample of cities in each period (**Table 3**). The main city transformations in this paper are: $I \rightarrow I$, $I \rightarrow II$, $II \rightarrow I$, $II \rightarrow II$. In **Table 3**, the magnitude of the diagonal values of the Markov matrix reflect the likelihood of cities maintaining their original scale class, while the non-diagonal values reflect the likelihood of interconversion between cities of different scale classes.

From **Table 3**, it can be concluded that: First of all, nearly 6.1%, 8.6%, and 3% of megalopolises grew into megacities in 2005-2010, 2010-2015, and 2015-2018, respectively, and the diagonal values of the matrix are all higher than 0.85, which reveals that the development of the scale of each class of cities is relatively stable, and the possibility of maintaining the original class is greater, and the scale of cities is not easy to transform. And then, the city class transition is mainly based on the transfer between cities of the same class. Among them, the mobility to the adjacent higher level is weak, in which the probability of transferring from type I to type II are all less than 0.1, and there is no reverse transfer from type II to type I. By this, it can be concluded that the probability of maintaining the original level of each city in the process of development transition is relatively high, and the possibility of city level transition is not significant. Consistent with the previous conclusion that the evolution of the city scale class has always been characterized by a concentration of city population towards megacities.

The mobility measure index of Markov transfer matrix (**Figure 8**) is calculated by Equation (6) and Equation (7), and then combined with the distribution and analysis of the city class scale in **Table 1**.

In the beginning, the overall level of liquidity tends to increase, with an increase (0.07) in the SMI from 2015 to 2018 (0.161) compared to the SMI from 2005-2010 (0.091). Specifically, the possibility of inter-transformation between the classes of cities shows that the diagonal element of class II is 1 from 2005 to 2015, indicating that the highest class of cities, *i.e.* (megacities), is extremely

Table 3. Markov matrix of city scale distribution in China.

particular year	scale	I	II	exit
2005-2010	I	0.909	0.061	0.03
	II	0	1	0
	enter	0.151	0	exit
2010-2015	I	0.857	0.086	0.03
	II	0	1	0
	enter	0.086	0.167	exit
2015-2018	I	0.939	0.030	0.03
	II	0	0.9	0.1
	enter	0.06	0.1	

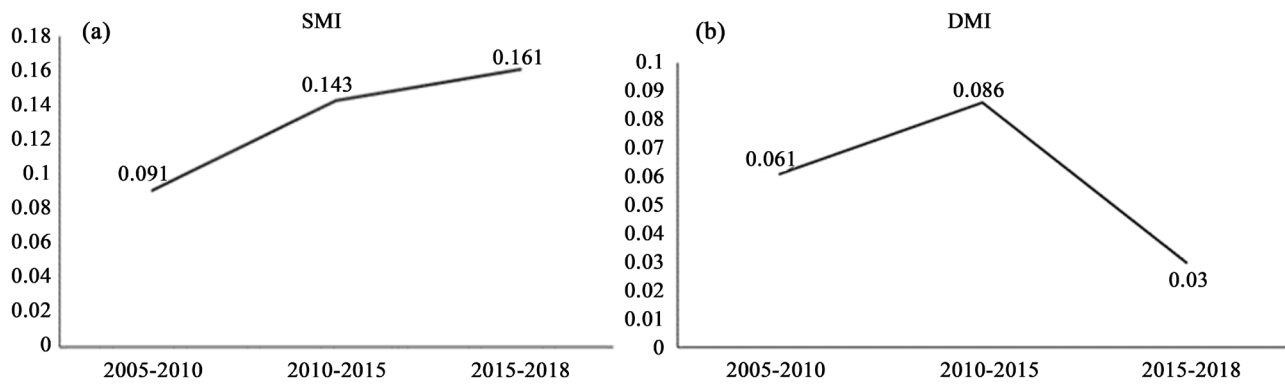


Figure 8. Markov matrix mobility trend of city scale distribution.

stable during this period, while the diagonal element of class I decreases, which implies that cities of class I scale become more mobile. Then, the diagonal element of class I becomes smaller from 2010 to 2015, illustrating those megalopolises become more mobile. Combined with **Table 1**, in this period, Yancheng, Cangzhou dropped out of the sample for the reason that Yancheng is not in the Yangtze River Delta economic circle, the economic base is weak, and the infrastructure is backward and the traffic is lagging behind, which seriously restricts the economic development, while the city area of Cangzhou is too small and far away from other cities, and the economic development is limited. Yueyang is a city with a developed oil industry at present, and the upgrading and optimization of the industrial industry has led to a focus on supporting the development of the city. Xiangyang was the second largest transportation hub in Hubei, with convenient transportation and gradually improving social and public services, so that both the scale of the economy and the population were increased and became a megalopolis at the same time as Yueyang during this period. The diagonal elements of rank II became smaller between 2015 and 2018, while the diagonal elements of rank I became larger, indicating that the megacities became more mobile at this time, combined with the fact that a new city (Zhengzhou) became a megacity in the previous section. However, Shijiazhuang's over-dependence on resources from Beijing and Tianjin led to over-capacity and stagnation in the development of the city's scale, and it eventually dropped out of the sample because its GDP per capita was lower than the national GDP per capita. The megalopolises became less mobile but remained stable in number during this period. Shaoxing is geographically adjacent to Hangzhou and Ningbo, and the implementation of the strategy of integrating Hangzhou with Ningbo and Shanghai to reach a deep connection with Shanghai has attracted a large number of talents, making the city scale develop rapidly in this period to become a megapolis. Zhangzhou is located on the west coast of the Strait, the good ecological environment which allows it to maintain a high economic growth rate, so that the population grows faster and becomes a megalopolis in 2018.

Following that, comparing the DMI index obtained from the non-diagonal elements can better illustrate the directionality of city mobility. From 2005 to

2018, city development shows a trend of upward mobility (DMI are larger than zero), and the elements on the left side of the diagonal of each class are smaller than the right side, and the left side element is always zero, therefore, city scale development maintains the trend of upward transformation *i.e.*, there is no reverse transformation from type II to type I.

Finally, considering new entrants and exiting cities in the sample, it is found that 15.1% of new cities entered rank I from 2005 to 2010, while the rank exit rate was lower; to the extent that more than 10% of cities entered rank II (megacities) in both time periods since 2010 to 2018, there is, moreover, evidence of a different city growth pattern in the two phases.

5. Conclusions and Recommendations

5.1. Research Findings

Through a series of analyses on the characteristics and dynamic evolution of the distribution of city scale in China, the following main conclusions are finally obtained:

1) From the perspective of megacities distribution regions in China, policy regulations have largely determined the basic pattern of city scale evolution. Firstly, strategies such as Reform and Coastal Opening, together with the development of Pudong New Area, have all contributed to the concentration of population in the eastern coastal region. Secondly, the city population scale in Chengdu, Chongqing and Central Plains has developed faster under the influence of the National Western Development Strategy.

2) The Rank-Scale rule fits the data of Chinese cities well and conforms to their distribution characteristics, that is, the population distribution is concentrated in high rank-order cities and the number of low rank-order cities is small. Shanghai and Chongqing have played a central role in regional development, have an obvious effect of radiation on the surrounding city. At this stage, the distribution of city population scale in China has shown the trend of gradually increasing the most populous city, which means that the phenomenon of unreasonable city population scale distribution is becoming more serious.

3) From 2005 to 2018, the Gini index of economic scale of China's cities is significantly higher than the Gini coefficient of city population scale, and both population and economy are in a state of regional concentration, with a high level of economic development and significant economic agglomeration benefits in megacities.

4) Judging from the nuclear density curve, the population gap between different cities is decreasing year by year, but the gap with the population scale of the most populous cities is still large, and it still takes a long time to achieve the harmonization of China's city population scale.

5) The diagonal elements of Markov transfer matrix are all higher than 0.85, indicating that the development of cities in these two levels of scale is relatively smooth, with a high probability of maintaining the original level, and the mobil-

ity of scale levels is not strong. Among them, the possibility of transferring from type I to type II in the process of development transition is small, and there is no reverse transfer from type II to type I.

5.2. Suggestions for Measures

Based on the constructed explanatory framework of the dynamics mechanism of city scale evolution (Figure 9), this paper proposes specific measures in four aspects: government measures, economic development, spatial pattern, and social livelihoods, respectively:

1) As for the government departments, they should formulate the city scale development measures based on the principle that the development patterns are different for different regions. The government should introduce appropriate preferential policies to encourage all cities to vigorously develop advantageous industries and improve city economies. At the same time, allow large cities to drive smaller cities to develop their economies and implement inter-city staggered development. For small cities with small populations, the government should focus on implementing talent acquisition programs to increase the scale of the population. For regions where excessive development of core cities leads to resource concentration, the government should appropriately control their development speed.

2) From the aspect of economic development. The economic development of cities must rely on industries, therefore, cities with slow development should

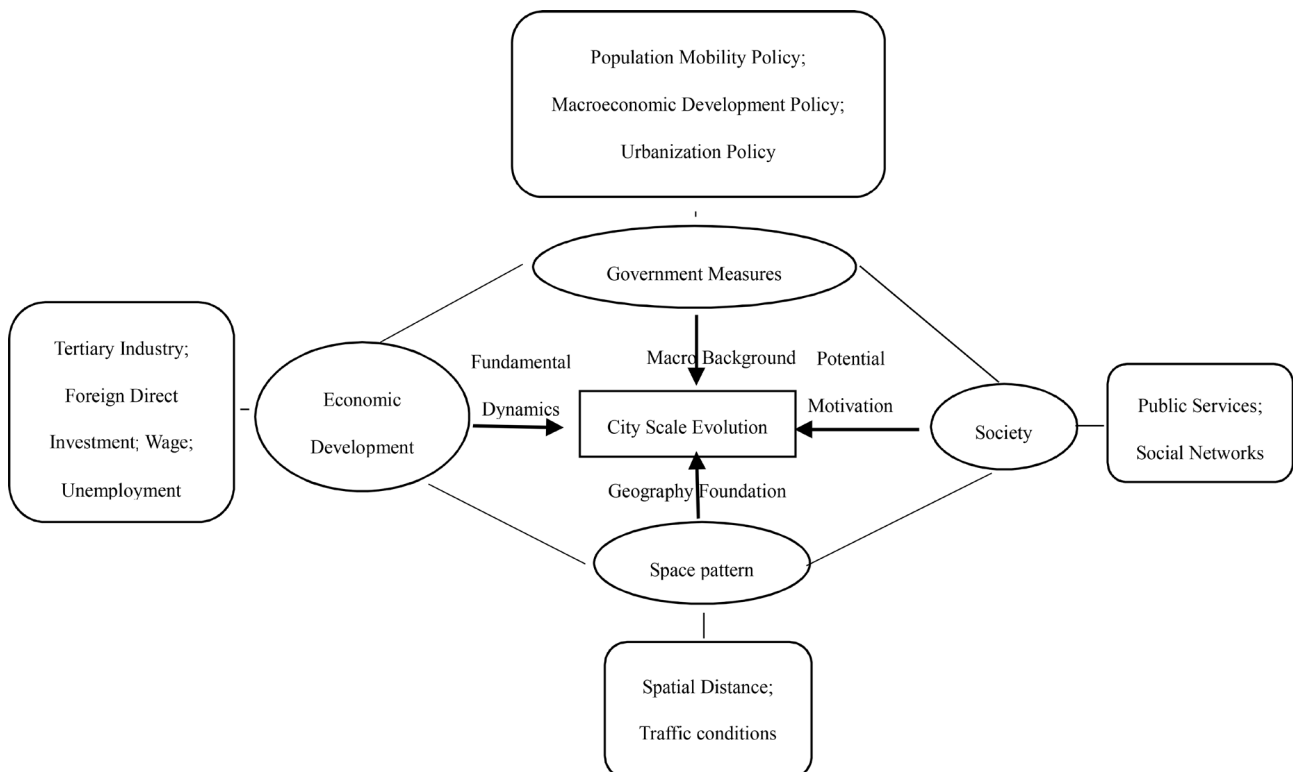


Figure 9. Theoretical explanatory framework of the dynamics mechanism of city scale evolution.

re-define their functional positioning and optimize the layout of industrial structure while vigorously developing special industries. Focus on cultivating secondary hubs in the region that can undertake industrial transfer in addition to the central cities and grow into new regional economic growth points, so that they can give full play to their role as bridges, allowing small cities to establish close economic ties with the central cities, promoting the flow of industrial capital between cities, realizing regional industrial upgrading and industrial transfer thus reducing city unemployment and raising city wage levels.

3) In terms of social livelihood, optimizing the service functions of the city, more employment opportunities, quality medical and educational resources, considerable income, and perfect service facilities are important factors to attract the inflow of middle and high-end talents. In addition, improving social network facilities and forming a safe cyberspace help cities develop.

4) The spatial pattern is the geographical basis of city development. The natural conditions and topographic features of our country lead to uneven distribution of population. Improving city transportation conditions, especially the external transportation construction of cities whose geographical location leads to obstructed communication with the outside world, improves transportation accessibility, reduces the spatial distance of cities, allows wider access to resources between cities, and thus promotes the development of cities with relatively weak economy.

Generally speaking, the samples in this paper were selected by setting thresholds based on the existing criteria for classifying city scale, however, different thresholds, as well as statistical, calibers can cause differences in the research results of city scale. In the future, we need to pay great attention to the problem of city sample selection and use multiple samples for comparative analysis as much as possible. In addition, attention should be paid to the changes in the division of the administrative areas of individual cities. Lastly, data sources should be expanded and more big data analysis tools should be used to conduct a more sophisticated study of the evolution of city scale.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Xu, X.Q., Monday, X. and Ning, Y.M. (2009) *City Geography*. 2nd Edition, Higher Education Press, Beijing.
- [2] Feng, J. and Zhong, Y.C. (2020) Spatial Structure of City Residents' Quality of Life in Changzhou Based on Living Environment. *Journal of Geography*, **75**, 1237-1255.
- [3] Jefferson, M. (1939) The Law of the Primate City. *Geographical Review*, **29**, 226-232. <https://doi.org/10.2307/209944>
- [4] Singer, H.W. (1936) The "Courbe des Population": A Parallel to Pareto's Law. *The Economic Journal*, **46**, 254-263. <https://doi.org/10.2307/2225228>

- [5] Marshall, J.U. (1997) Beyond the Rank-Scale Rule: A New Descriptive Model of City Scales. *City Geography*, **18**, 36-55. <https://doi.org/10.2747/0272-3638.18.1.36>
- [6] Feng, X.H., Zhong, Y.X., Li, Z.R., *et al.* (2017) Evolvement of Spatial Pattern of City System in the Economic Belt of Yangtze River. *Resources and Environment in the Yangtze Basin*, **26**, 1721-1733.
- [7] Wei, C.J., Zhu, J.G., Li, X.J., *et al.* (2017) A Rethink of the Rank-Scale Rule for Rural Settlement in Traditional Agricultural Areas: A Case Study of Zhoukou City. *Economic Geography*, **37**, 158-165.
- [8] Wan, Q., Wu, C.Q., Luo, X., *et al.* (2018) Spatial and Temporal Evolutionary Characteristics of City Scale Distribution in China: An Empirical Study Based on the Demographic Data of the “Five” and “Six” Censuses. *Economic Geography*, **38**, 81-90.
- [9] Li, S.J., Xu, J.W., Hai, B.B., *et al.* (2015) Analysis of the Evolution of County Settlement Distribution Pattern: An Empirical Study Based on Gongyi, Henan, 1929-2013. *Journal of Geography*, **12**, 1870-1883.
- [10] Li, S.S. and Zhang, Z.Y. (2015) The Impact of City Scale on City-Rural Income Gap—An Analysis Based on Panel Data of Prefecture-Level Cities. *City Issues*, **6**, 14-20.
- [11] Zhou, Y., Huang, X.J., Xu, G.L., *et al.* (2016) The Coupling Dynamics of City Land Expansion and Population Growth in the Yangtze River Delta and Its Driving Mechanism. *Geography Research*, **35**, 313-324.
- [12] Liu, Y.B., Wang, Y. and Xie, F. (2013) Characteristics of the Evolution of City Scale Structure in the Poyang Lake City Cluster. *Economic Geography*, **33**, 70-76.
- [13] Liu, B.H. and Liu, Z.L. (2009) Research on City Function Structure and Its Transformation in Central Liaoning City Cluster. *Economic Geography*, **29**, 1293-1297.
- [14] Zipf, G.K. (1949) *Human Behavior and the Principle of Least Effort*. Addison-Wesley Press, London.
- [15] Xu, Y.T., Liu, X.Z. and Wang, Z.B. (2018) Spatial and Temporal Distribution Characteristics of City Air Quality in China Based on AQI Index. *Journal of Guangxi Normal University (Natural Science Edition)*, **37**, 187-196.
- [16] Wang, Z.P., Zhang, X.F. and Wang, L.Y. (2018) Study on the Distribution of City scale and Evolution Mechanism of Zhongyuan City Agglomeration. *Regional Economic Review*, **1**, 108-116.
- [17] Jiang, Y.L. and Zhang, D. (2018) Study on the Change of City Scale Distribution and Spatial Characteristics of the Beibu Gulf City Cluster. *Regional Governance*, **3**, 33-34.
- [18] Wan, Q., Wu, C.Q. and Feng, B. (2019) Spatio-Temporal Evolutionary Characteristics of City Scale Distribution in China—An Empirical Analysis Based on DMSP/OLS Nighttime Lighting Data. *Regional Economic Review*, **4**, 125-133.
- [19] Cao, Y.C. and Zhang, Y.K. (2015) Study on the Scale Structure and Anisotropic Growth Relationship of City System in Hebei Province. *Arid Zone Resources and Environment*, **29**, 13-18.
- [20] Shorrocks, A. (1978) The Measurement of Mobility. *Journal of the Econometric Society*, **46**, 1013-1024. <https://doi.org/10.2307/1911433>
- [21] Eaton, J. and Eckstein, Z. (1997) Cities and Growth: Theory and Evidence from France and Japan. *Regional Science and Urban Economics*, **27**, 443-474. [https://doi.org/10.1016/S0166-0462\(97\)80005-1](https://doi.org/10.1016/S0166-0462(97)80005-1)