

Zipf Law Analysis of Urban Scale in China

Wen Ziqin¹

¹ Graduate School of education, Jinan University, Guangzhou, China

Correspondence: Li Yang, School of economics, Jinan University, Guangzhou, China. Tel: 86-155-2133-8550.

E-mail: kobe0823yang1022@163.com

Received: February 1, 2016

Accepted: February 5, 2016

Online Published: February 29, 2016

doi:10.20849/ajsss.v1i1.21

URL: <http://dx.doi.org/10.20849/ajsss.v1i1.21>

Abstract

In this paper, by using China's urban population data from 1990 to 2010 and double logarithmic regression model to test China's urban scale and urban rank through Zipf law, we found that China's urban scale distribution is relatively balanced, virtually conforming to Zipf's law. Researching and analyzing urban scale distribution of China's every province's, we also found that the urban scale distributions of China's all provinces are relatively complicated and they are not even close to 1. The increase of Zipf index means that urban scale distribution is transiting from primacy degree distribution to ideal Zipf distribution. Further analyzing and categorizing every province's Zipf index, we can conclude that China's urban scale distribution can be categorized four mainly: achieved ideal Zipf law; approaching to ideal Zipf distribution; The transition state from the primacy degree distribution to the Zipf law distribution; typical primacy degree distribution.

Keywords: Zipf law, urban distribution, urban scale

1. Introduction

In the research of city size distribution, the urban scale and the urban level have good characteristics. The earliest research on urban scale and order is from Auerbach (1913) and Singer (1936). They believe that the scale of the city and the sequence can be characterized by Pareto distribution, and there is a mathematical relationship of:

$y = A \cdot x^{-\alpha}$. Zipf (1949) believes that the scale of the city can not only be expressed by the Pareto distribution,

and when $\alpha = 1$, it found $R_i \cdot S_i = A$, Zipf index, the relationship between the size of the city and the order of

the city is called the Zipf rule. Since 1950s, a large number of empirical studies have been carried out on the Zipf law. Berry (1961) by the 38 countries on the urban population data for empirical testing, the city size distribution is divided into three categories: Firstly, including 13 countries, in full compliance with the size of the city and the sequence distribution, that is, the city size distribution of the Zipf rule. Secondly, including 15 countries, including the first city to control the scale of urban structure characteristics; thirdly, there are 10 countries balance between the above characteristics. Madden (Madden) using the United States city population data from 1790 to 1950, the use of a logarithmic model to test the stability of the Zipf rule. The conclusion shows that the urban structure system changes very quickly in the scale, and the city's rank is also found changing, but the city size distribution itself is stable in a long time. Soo (2004) using 75 national urban population data, using the method of least squares fitting and maximum likelihood estimation to compare the Zipf index of several countries in the world. The results show that the distribution of the countries in the world is basically in accordance with the Zipf law, and the law is stable. The domestic scholar Xu Xueqiang (1993) tracked the urban scale distribution of the top 1953 in China in the 1960 - 1990, and predicted the situation in 2000.

Through the Zipf rule test of the city size distribution in China, the characteristics and the changing trend of the city size distribution in China are analyzed. The study is divided into four parts. First part is a literature review; the second part is the part of empirical research, mainly on the model set and research data do a brief description;

the third part of the empirical results were analyzed; the fourth part is a brief conclusion of the research.

2. Model and Data

2.1 Model Settings

We mainly uses the following model to test the size distribution of cities in China:

$$\ln R_i = \ln A - \alpha \ln S_i \quad (1)$$

Where R is the city level, S is the size of the city, α is the Zipf index, the Pareto value, but also need to estimate the data equation. The main purpose of the study is to study the size and change trend of the Zipf index of urban size distribution in China.

In the city size distribution, the Zipf index (α value) shows the aggregation of the city size distribution. The natural logarithm of city scale and registration is calculated, and the regression of these data is obtained. The significant log linear model is the order - scale rule. If the slope of the regression is -1 ($\alpha = 1$), then the city size of the Zipf rule can be verified. If $\alpha > 1$, shows that the city's population is relatively scattered, small and medium-sized city scale development is good, the high level of urban size is not prominent; if $\alpha < 1$, proved that the scale of the distribution is concentrated, small and medium-sized urban population size is poor, the size of large urban population growth, urban size distribution of the first degree is higher. In Zipf's index to study the trend of changes, α in becomes large, show that the city size distribution tends to be concentrated strength is greater than that of the dispersion of power ;if α is smaller, indicating that makes the city scale distribution tends to disperse the strength is greater than the concentration of power.

Gabaix and Ioannides (1975) proposed using Hill method to replace the least squares method to calculate the distribution of urban size Zipf index. The Hill method is essentially a maximum likelihood estimation method in the case of the model. If the sample size is n , the expression of the size of the city is: $x_1 > x_2 > x_3 > \dots > x_n$,

The city size distribution of Zipf index can be calculated by the following equation:

$$\alpha = \frac{n-1}{\sum_{i=1}^{n-1} (\ln x_i - \ln x_n)}$$

Standard deviation representation as:

$$\sigma_n(\alpha) = \alpha^2 \left\{ \frac{\sum_{i=1}^{n-1} (\ln x_i - \ln x_{i+1})^2}{n-1} - \frac{1}{\alpha^2} \right\}^{1/2} n^{-1/2}$$

Through the empirical test of the Zipf rule of China's city size distribution, the ordinary least square method is adopted. The least square error of the least square method is relatively small, and it can fit the relationship between the size and the position of the city. The premise of the Hill method is that if the city size obeys Pareto distribution, the Zipf index of urban scale is tested, which is not representative and general.

2.2 Data Description

An urban scale is the population, economy and science and technology in a certain area of the size of the aggregation. The general scale of the city includes population scale, land use, building and facilities, and the scale of productivity and consumption. Narrow city scale only refers to the population size in a certain city. Compared to other measures of city size, the relevant data of urban population size is more easy to collect, but also the most commonly used indicators. The non-agricultural population in China's urban residence is the main measure of the urban scale. Take non agriculture population is: first, non-agricultural population in number and the built area of the resident population is close to; second, for the urban system, each city non-agricultural population and built district resident population in general showed a linear relationship; third, non-agriculture in

some extent to maintain statistical consistency and comparability.

Secondly, because of Chongqing, Beijing, Shanghai, Tianjin and other four municipalities have their own special characteristics of the scale of the city, the return of the four cities are not considered. In addition, Macau and Hong Kong in the late 1990s to review the motherland, so Macao and Hong Kong's urban population data exist flaw phenomenon is serious, so we in the establishment of measurement model, without considering the special situation of Macau and Hong Kong. Similarly, for the city of Taiwan scale, also temporarily do not consider. Finally, due to the vast territory of China, the statistical data of the urban population of Tibet autonomous region has a serious deficiency in the statistical yearbook of China in recent years. And Tibet have a certain number of the scale of the city is relatively small, in order to prevent the statistical results and the model brought large errors, we selected China statistical yearbook data of city size distribution test will Tibet Autonomous Region exclude outside.

According to the original data of the 2010 - 1990 China Urban Statistical Yearbook, a descriptive statistical table of the urban population was obtained, which was shown in Table 1.

Table 1. Descriptive statistics of urban population in China from 1990 to 2010

Unit: 10 thousand population

Year	1990	1995	2000	2005	2010
Sample quantity	467	640	647	655	655
Min	0.31	0.51	0.30	1.58	1.63
Max	749.56	833.80	938.21	938.20	938.24
Med	16.16	17.05	17.36	17.36	17.42
Average	32.20	31.28	33.23	33.33	33.36

3. Empirical Result

3.1 The Results of Regression Analysis of China's Urban Scale Distribution from 1990 to 2010

According to regression equation $\ln R_i = \ln A - \alpha \ln S_i$, based on Matlab R2013a, by using least square fitting to analyze China's population from 1990 to 2010 with regression model, the results and analysis are shown in Table 2.

Table 2. The results of regression analysis of china's urban scale distribution from 1990 to 2010

	Coefficient	Standard Deviation	T Value	PValue
1990	-0.9031	0.01844	-37.7121	0.0000
1995	-1.0231	0.01526	-36.5393	0.0000
2000	-1.0405	0.01325	-46.6093	0.0000
2005	-1.0986	0.01435	-46.9892	0.0000
2010	-1.0426	0.01237	-43.4343	0.0000

According to the results from Table 2, P value are all 0 when using double logarithmic model to do regression analysis. They can all pass the hypothesis test with confidence level of 1%. Thus, we can conclude that the regression analysis model is significant. The model well explains the good linear relationship between the urban scale and the order of the city. To better describe the trend of Zipf index of urban scale distribution from 1990 to 2010, we made Figure 1 based on Table 2.

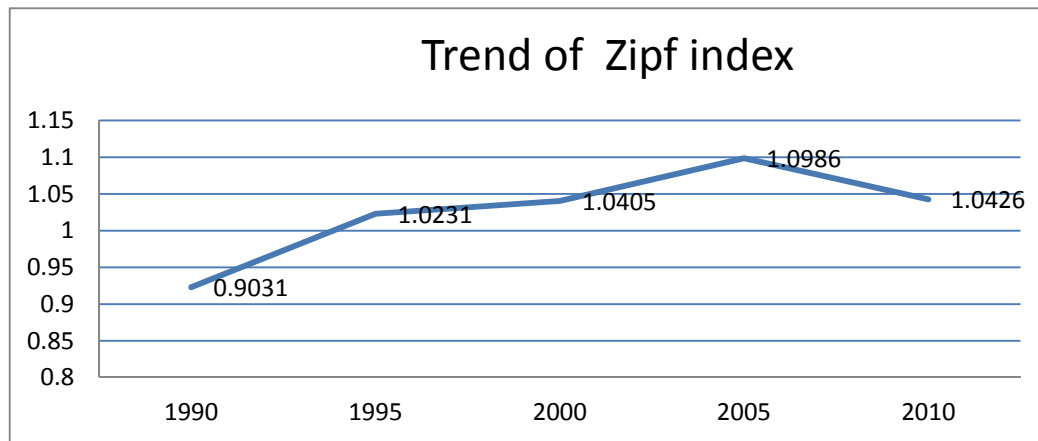


Figure 1. The trend of Zipf index of urban scale distribution from 1990 to 2010

It is shown that there are minor fluctuation of Zipf index of urban scale distribution around $\alpha = 1$, which means that Zipf index of China's urban scale distribution from 1990 to 2010 are balanced, conforming to the ideal status of Zipf distribution. Its basic trend is that Zipf index increased steadily ifrom 1990 to 2005 but it began to fall around 2010. The urban scale distribution index in 1990 was $\alpha = 0.9031 (\alpha < 1)$. This shows that China's urban population distribution is relatively concentrated in 1990, the primacy degree is higher, the larger cities are relatively prominent, while the small and medium sized cities are insufficient. Compared to 1990, urban scale distribution in 1995 and 2000 were more close to 1, which means that urban scale distribution were more balanced and reached the ideal status of Zipf distribution. Zipf distribution of 2005 was relatively high, however it fell back to ideal status in 2010.

3.2 Testing Analysis of China's Urban Scale Provinces from 1990 to 2010

From 1990 to 2010, China's urban size distribution equilibrium, close to the ideal state of the Zipf index distribution. At the same time, in recent years, China's economy has also been a good development, the city's size distribution has also changed significantly. To this end, in order to explore the status of the distribution of urban size and change in various provinces and cities, by the size of cities in various provinces of China Zipf rule test.

Analysis table 3 can found that, in the 1990 - 2010s, the Zipf distribution of the provinces in China is more complex, and not significant. Between the provinces, and of different periods, the distribution of urban size differences are great. Further analysis found that in 1990 China's urban size distribution in the Zipf ideal state of the province a total of 13 ($0.8 < \alpha < 1.2$), which shows that the size distribution of cities in various provinces of China has not reached the ideal state distribution. And in the 1995, 2000, 2005, 2010, Avenue equilibrium distribution of provinces respectively, 19, 19, 26, 26, from a nationwide perspective, Chinese city scale in provinces in recent ten years of the Zipf distribution index in general has reached the ideal state of the Zipf distribution.

In order to further study the China in 1990 - 2010 years all the provinces of the urban scale distribution, we're on the regression analysis of Zipf's index evolution of similarity analysis, to classify the various provinces of the country of urban scale. The results shows that Chinese provinces, city size distribution is mainly divided into the following categories:

The first category is Zipf's index is relatively high, indicating that the provinces and cities development is good, and the first city size on the influence of the size of the population of the province's is not obvious, the city size and rank are more balanced distribution. Including Liaoning, Hunan, Sichuan, Xinjiang, Henan, Anhui.

The second category are the provinces city size distribution is basically consistent with the Zipf's law of an ideal state of distribution, provinces with the central city, and small and medium-sized cities scale development of good, high order cities complete is generally China's economy is relatively developed, or scale economy relatively large provinces. Including Guangdong, Zhejiang, Jiangsu, Jilin, Hubei, Shandong, Fujian.

The third category is the Zipf's index is relatively small, belonging to the general economy is relatively less

developed provinces, area economy is relatively backward, center city development is not perfect, resulting in center city scale, thus highlighting has obvious first features. At the same time, the rapid development of economy in recent years promoted the rapid development, two provinces city, province is in the first degree distribution to Zipf ideal equilibrium distribution of the transition state, these provinces including Guangxi, Jiangxi, Shanxi, Hebei, Shaanxi, Guizhou, Gansu.

The fourth category is the Zipf's index is very small, provinces in general in addition to the center of the capital city, other cities are not perfect development, urban population highly concentrated in the capital city, is typical of the first feature, the city size distribution belongs to the first distribution. Including Ningxia, Qinghai, Inner Mongolia, Yunnan, Heilongjiang.

Table 2. Zipf index measurement regression results of urban scale distribution in provinces of 1990-2010

	1990	1995	2000	2005	2010
Hebei	0.640	0.764	0.774	0.828	0.801
Henan	0.947	1.040	0.982	0.970	0.976
Heilongjiang	0.567	0.656	0.832	0.981	0.9065
Hubei	1.032	1.079	1.133	1.148	1.1405
Hunan	0.853	0.896	0.863	0.901	0.882
Jilin	1.074	1.190	1.173	1.213	1.193
Jiangsu	1.026	1.182	1.185	1.067	1.126
Jiangxi	0.864	1.056	1.035	1.035	1.035
Liaoning	0.854	0.872	0.833	0.869	0.851
Inner Mongolia	0.634	0.686	0.738	0.809	0.7735
Ningxia	0.577	0.599	0.736	0.898	0.817
Qinghai	0.357	0.375	0.398	0.376	0.387
Shandong	1.905	1.186	0.996	1.206	1.101
Shanxi	0.767	0.839	0.885	0.896	0.8905
Shanxi	0.738	0.786	0.797	0.716	0.7565
Sichuan	0.968	1.028	1.134	0.919	1.0265
Xinjiang	0.894	0.938	0.981	0.926	0.9535
Yunnan	0.440	0.5788	0.889	0.903	0.896
Zhejiang	1.052	1.192	1.193	0.996	1.0945
Anhui	1.026	1.109	1.078	0.974	1.026
Fujian	0.953	1.158	0.969	1.109	1.039
Gansu	1.766	0.816	0.880	0.802	0.841
Guangdong	1.045	1.441	1.126	0.844	0.985
Guangxi	0.664	0.870	0.528	0.983	0.7555
Guizhou	0.739	0.810	0.790	0.836	0.813

4. Conclusion

Through the collection of urban population data from 1990 to 2010, the Zipf rules of urban size and urban order

were tested and found that the size distribution of Chinese cities was more balanced, which was basically consistent with the Zipf rule. Zipf index values from 1990 to 2005 continued to grow, showing an increasing trend in 2010 Zipf index fell more close to the Zipf distribution of the ideal state.

Through the research and analysis of the city scale data of each province in China from 1990 to 2010, the city size distribution of each province in China is more complex, and which is not 1. Zipf index gradually increased, indicating that the city size distribution gradually changed from the first degree distribution to the Zipf rule. At the same time, for all provinces of city size distribution of Zipf's index were further analysis classified, found across Chinese provinces city size distribution state can be divided into four types: already reached the ideal state of Zipf's law distribution; close to the ideal of Zipf's law distribution; by the first degree distribution to the Zipf's distribution of the transition state; typical first degree distribution.

Reference

- Berry, B. J. (1961). City size distribution and economic development. *Economic Development and Cultural Change*, 9, 573-587. <http://dx.doi.org/10.1086/449923>
- Chen, Z. (2008). A review of theory and empirical research on optimal city size. *International Urban Planning in China*, 6, 87-92. <http://dx.doi.org/10.3969/j.issn.1673-9493.2008.06.012>
- Deng, Y., Liu, S., Cai, J., & Lan, X. (2014). Comparison of the methods of urban influence scope division in Central China. *Chinese Geographic Research*, 7, 77-82.
- Madden, C. J. (1956). Some indications of stability in the growth of cities in the United States. *Economic Development and Cultural Change*, 4, 236-452. <http://dx.doi.org/10.1086/449715>
- Soo, K. T. (2005). Zipf's Law for cities: a cross-country investigation. *Regional Science and Urban Economics*, 35(3), 239-263. <http://dx.doi.org/10.1016/j.regsciurbeco.2004.04.004>
- Tian, J. (2014). Review of Zipf's law of city size distribution. *Science and Industry in China*, 7, 34-39. <http://dx.doi.org/10.3969/j.issn.1671-1807.2010.07.010>
- Xu, X., & Ye, J. (1996). Inter provincial differences of urbanization in China. *Journal of Geography in China*, 7, 55-62. <http://dx.doi.org/10.3321/j.issn:0375-5444.1986.01.002>
- Ye, H., Pu, L., & Zhang, P. (2012). The inter provincial differences of urban spatial distribution in China and its influencing factors. *China's population resources and environment*, 12, 65-71. <http://dx.doi.org/10.3969/j.issn.1002-2104.2012.12.025>
- Yu, J., Zhou, G., & Duan, Y. (2013). Research on the evolution trend of urban scale distribution in China -- Based on the National Population Census. *Population and Economics in China*, 2, 43-50. <http://dx.doi.org/10.3969/j.issn.1000-4149.2013.02.006>
- Zipf, G. K. (1949). *Human Behavior and the principle of Least Effort*. MA: Addison-Wesley Press.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).