CITY SIZE DISTRIBUTION, CITY GROWTH AND URBANISATION IN CHINA

By

YAN CEN

Thesis submitted to The University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

Department of Economics

The University of Birmingham

April 2015

ABSTRACT

This thesis explores three topics within the broad area of urban growth and environment in China, city size distribution, city growth pattern and the environment impacts of city growth. The research is firstly motivated by two key stylized facts- the well-known Zipf's law for cities (which states that the number of cities of size greater than S is proportional to 1/S, i.e. the rank of a city is inversely correlated with its size) and Gibrat's law for cities (which states that city growth rate is independent of its size). Thus Chapter 3 and 4 examine the evolution of city size distribution by testing for Zipf's law and Gibrat's law in China from 1879 to 2009 (number of cities varies over time). Chapter 5 thereafter investigates the growth pattern of Chinese cities by testing for the sequential city growth (Cuberes, 2009). Given the concern of the environment impacts of city growth, Chapter 6 examine the impact of city size on local air quality using 30 major cities in China from 2003 to 2012.

Our findings suggest that Zipf's law emerged in the end of 1980s and early 1990s, before that Chinese city size distribution is less even than Zipf's law predicts while after that the city size distribution is more even than Zipf's law predicts. Gibrat's law involves as an explanation of Zipf's law (Gabaix, 1999), results confirm that Chinese cities might follow Gibrat's mode of urban growth- stochastically growth under a homogeneous urban growth process with a common mean and variance. Secondly, with respect to the urban growth pattern, results show that Chinese cities tend to grow sequentially, with the largest city grew the fastest initially and then the second largest city grew the fastest, if we use the same econometric method as Cuberes (2011). Thirdly, the results of the impact of city size on local air quality are mixed (details show in Chapter6- results table) and also extend the EKC hypothesis to city total GDP level, i.e. city total GDP or economic scale has an inverted-U shape relationship with pollution. The results from this thesis both contribute to the literature on urban growth and environment and the urban policy makers in terms of designing efficient urban development policy.

Key words: city size distribution, Zipf's law, Gibrat's law, sequential city growth, air quality

ACKNOWLEDGEMENTS

My thanks go primarily to my supervisors Professor Matthew Cole and Professor Robert Elliott who are the best supervisors in the world. I would like to appreciate them for being patient always and providing invaluable guidance and encouragement throughout my time at Birmingham.

I am also grateful to the Department of Economics in University of Birmingham for providing the bursary funding for my PhD. Many thanks also go to members of the department and research students for all their assistance.

Finally, special thanks go to my parents and beloved sister Chen, for their love, constant support, patience, and encouragement during the writing of this thesis.

TABLE OF CONTENTS

CHAPTER 1
INTRODUCTION

An important development in the global economy is that economic activity is increasingly undertaken in cities with Lucas (1988) pointing out that cities are the primary engines of economic growth. Such a finding is not surprising given the rate of urban growth or urbanization that has occurred over the last century. For example, in 1950, 30% of the world's population lived in cities. In 2000 this fraction grew to 47% and is predicted to rise to 60% by 2030.[1] In the U.S., 80% of the population live in urban agglomerations and they earn around 85% of income (Rossi-Hansberg and Wright 2007). Therefore, issues related to patterns of urbanisation and city growth have attracted the interest of academics and policymakers.

The impact of increased urbanisation is particularly important in the case of China which has experienced a remarkable transformation from a largely rural and agricultural economy to a modern manufacturing centre. This economic transformation has taken place at the same time as huge numbers of Chinese workers have moved from the rural areas into existing and newly built cities. It is this growth in urbanization and the unique government structure in China (and changes in government policies) that were the motivation for this thesis.

Our focus on China is because China is currently the largest developing economy in the world and has undertaken remarkably rapid urbanisation. Nearly 0.666 billion

[1] United Nations, " World Population Prospects: The 2004 Revision Population Database" (esa.un.org/unpp [October 2005])

Chinese people live in cities (including county-level cities), which means that nearly half of the Chinese population is currently living in urban areas ('Sixth Census of China 2010' released in April 2011).

Despite the rapid growth of the last three decades, new construction and urban sprawl continues. Against a background of rapid urbanisation, it is important for poliymakers and acdemics to understand how cities develop and how their distribution changes over time. Firstly, a knowledge of city size distribution and growth patterns will help policymakers make informed decisions on, for example, urban infrastructure, education, health investment and environment management. For example in Chapter 6 we link city size and air quality pollution which may provide advice for urban planners on how city size can affect local air quality. Secondly, our study of city size distribution, city growth pattern can also contribute to the acdemic study of the stylized empirical regularity- 'Zipf's law' and 'Gibrat's law' and sequential city growth (Cuberes, 2009).

The thesis consists for four substantive chapters (Chapter 3, 4, 5, 6). After a review of the data and Chinese urban system in the second chapter, this thesis is initially motivated by testing for the two stylized facts of city size distribution- Zipf's law and Gibrat's law (chapter 3 and chapter 4). Zipf's law refers to the regularity that within a country, the number of cities of size greater than S is proportional to 1/S, i.e. the rank of a city is inversely correlated with its size and the power is -1. In other words, within a country, if one rank the city by their population size decreasingly, then one can observe the fact that 1st ranked city roughly has twice the population of 2nd ranked city, and has roughly three times the population of 3rd ranked city, etc. With respect to Gibrat's law, it refers to that within a country, city growth rate is independent of city size (population size), i.e. bigger cities may not grow faster than smaller ones. The fifth chapter studies the pattern of growth of Chinese cities by testing for the sequential city growth theory. Finally, in related research we examine an area of importance to academics and policymakers to investigate the environmental consequences of urban growth by exploring the relationship between city growth and local air quality conditions.

Each of the main chapters in this thesis can be read independently but they are also internally linked and have a coherent logical sequence. This thesis is primarily empirical in nature but all chapters are based on theoretical models that have been studied in the existing literature.

Specifically, in Chapter 3, Zipf's law was proposed by American linguist George Kingsley Zipf[2] in 1949, which states that within a country, city rank (rank by

[2] Zipf's law was firstly used to address the fact that the frequency of any word is inversely proportional to its rank in the frequency table [2](Zipf 1935, 1949). This relationship has subsequently been found in

population) is perfectly inversely correlated with city size (population size), i.e. 1st ranked city roughly has twice the population of 2nd ranked city, and approximately has three times the population of 3rd ranked city, etc.[3] Zipf's law for cities has been proved to hold empirically both across countries and over years (Rosen and Resnick 1980; Krugman, 1996; Eaton and Eckstein, 1997; Gabaix, 1999a; Dobkins and Ioannides, 2000; Davis and Weinstein, 2001; Ioannides and Overman, 2003; Soo, 2005; Rose, 2005). Researches supporting Zipf's law mainly focus on developed countries and relatively large cities. However, some researchers find that Zipf's law does not hold well using different testing methods or different countries or time periods (Black and Henderson, 2003; Eeckhout, 2004[4]; Garmestani *et al.*, 2007; Soo, 2014). Therefore, we explore the question of whether there is a universal law for the size distribution of cities and whether the Chinese city size distribution follows Zipf's law.

In order to study the evolution of the city size distribution we use Chinese city-level population data from the end of Qing Dynasty in 1879 to the modern period with data up to 2009 (the number of cities for each year varies from around 30 to over around 600). Our findings suggest that over time the Chinese city size distribution has become more even (equal) than Zipf's law would predict, i.e. the disparity between

---

many other areas like physical and social sciences such as the population ranks of cities in various countries, corporation sizes, income rankings and so on.

[3] More detailed derivation of Zipf's law can be found in Chapter 3.

[4] Eeckhout (2004) proves that if one considers all U.S. places then city size distribution follows a log-normal distribution and not Pareto distribution (including Zipf's law).

large and small cities is smaller than Zipf's law would predict over time. However, we find that Zipf's law approximately holds in the end of 1980s and early 1990s. In addition, we divide our data into prefecture-level cities and county-level cities; four economic regions (East, Central, West, and North-East cities); and four historical city groups. Results for these subsamples are mixed.

In Chapter 4 we test Gibrat's law for Chinese cities. Gibrat's law originally argues that if the size of a firm (or a fund or a city) is S, then the growth rate of the firm (or a fund or a city) is independent of its size S and was proposed by Gibrat in 1931. Gibrat's law is firstly introduced into the Zipf's law debate as the explanation of Zipf's law, Gabaix (1999). Gabaix proposed that if various cities grow at the same mean and variance then the city size distribution will tend towards Zipf's law. Therefore, we test whether the mode of Chinese city growth (population growth) is a random process, and whether city's growth rate is correlated with city size, i.e. whether larger cities will grow faster or whether small cities will grow slower.

We use the same dataset that we use in Chapter 3 and our findings suggest that Gibrat's law gradually emerged although it was not fully attained, which could explain the results for Zipf's distribution in Chapter 3. However, unlike Gabaix (1999) 's argument which states that Zipf's law will emerge as the steady state, in our study the Chinese city size distribution firstly shows less evenness than Zipf's law predicts, and then it is consistent with Zipf's law predictions for a few years (end of 1990s and early 2000s), finally it becomes more even (equal) than Zipf's prediction, i.e. the disparity between large and small cities is increasingly smaller during 1879 to 2009 and finally grows beyond Zipf's law prediction.

The contribution of the first two main chapters 3 and 4 is twofold. Firstly, we construct a unique dataset that, for the first time, allows us to investigate the evolution of Chinese city size distribution and the growth pattern over a century.[5] Secondly, for the first time for Chinese cities in city size distribution studies, we divide the whole sample into several subsamples according to their administrative attributes, economic region, and historical experiences, in order to investigate the size distribution in more comprehensive perspectives.

In Chapter 5 we extend Chapter 3 and 4 which study spatial agglomeration in a static context, i.e. both Zipf's law and Gibrat's law describe the city size distribution at a given point of time. In Chapter 5 we extend our Chinese urban analysis to a dynamic model- - whether Chinese city growth follows the Sequential city growth theory first looked at by Cuberes (2011). The theory claims that within a country the largest city will grow the fastest initially, and then as time passes the second largest city will

[5] From 1984 to 2009 data are provided by the 'China Urban Statistical Yearbooks' (National Bureau of Statistics of China, 1985-2010). Before 1983 back to the 1879, data is provided by Jan Lahmeyer (http://www.populstat.info/)

become the ‘fastest grower’, and then at some point eventually the third-largest city starts growing the fastest and so on. Therefore, we investigate whether Chinese cities grow sequentially. In addition we assume the growth sequence is according to city size implicitly in Cuberes (2009)'s model. However, we also test whether there is sequential growth based on a city’s age (Sanchez-Vidal *et al*., 2014).

We use the same data for urban population as we used in Chapters 3 and 4.[6] Our findings suggest that if we use the same econometric methods employed in Cuberes (2009), then Chinese cities' growth pattern tends to follow a sequential city growth pattern. With respect to the age sequential city growth our results show that in China, unlike the Sanchez-Vidal *et al*. (2014)'s results showing that young cities tend to grow the fastest and then the growth rate slows down as time passes, for China older Chinese cities tend to grow the fastest initially, and then slows down. This chapter contributes to the empirical evidence of sequential city growth of Cuberes (2009) and Cuberes (2011).

Finally, in Chapter 6 we consider the concepts raised in the previous chapters on city size and city growth to issues related to the city environment. This is motivated by the obviously increasing urban environmental problems in China. Outdoor air pollution contributed to 1.2 million premature deaths in China in 2010, nearly 40% of the

[6] We update 2010 population data in Chapter 5 and the number of cities varies over year.

global total.[7] Therefore, we study the impact of city size on local air quality using panel data for 30 major cities (province capitals) in China for 10 years from 2003 to 2012. Specifically, we employ four indicators to represent city size: city population size, urban area size, total GDP (economic) size, and energy consumption size. We then use three air pollutants as alternative dependent variables (PM_{10}, SO_2, NO_2). Therefore, we are able to test firstly whether cities with large population sizes or urban area sizes have better or worse air quality, secondly, whether large cities in GDP terms have better or worse air quality and thirdly whether large energy consuming cities have better or worse air quality.

The contribution of this chapter is threefold: firstly, unlike the previous literature that mainly focuses on the economic growth and environment relationship, we directly address the impact of city size on the local environment and use various indicators to represent city size. Secondly we focus specially on China given the dramatic urbanisation process and given the lack of previous research in this area. Finally, in light of significant foreign direct investment into Chinese cities, we also consider the environmental impact of industrial output by ownership category and differentiate between output that is domestically owned, foreign owned and HMT owned (Hong Kong, Macao and Taiwan). Our findings for the three air pollutant are mixed.

---

[7]. According to Ministry of Environmental Protection of the People's Republic of China's report, 2010.

In the last Chapter 7, we summarise the whole thesis and address the results from previous empirical chapters. We also briefly discuss of potential future research.

CHAPTER 2
CHINESE URBAN DEVELOPMENT AND THE ADMINISTRATIVE DEFINITION OF CITIES

2.1 DATA SPECIFICATION

In China, cities are defined from an administrative perspective, although the precise definition of a city is an issue that is often debated in the literature. Some studies argue that using the administrative definition of cities may lead to biased results because even if the official statistics are reliable they are still based on the authorities' definition of city boundaries which may or may not coincide with the economically meaningful definition of "city" (see Rosen and Resnick, 1980 and Cheshire, 1999). It has been argued that agglomeration data more closely match the functional definition of city, and research has confirmed Zipf's law using this kind of agglomeration data or "area clustered" data (for a recent review see Rozenfeld *et al*., 2009). However, in the case of China, administrative cities are likely to be very similar to cities defined in terms of economic population clusters. Due to the long history of human settlements starting from thousands of years ago[8], the 'city proper population' data, on some level, is indeed the agglomeration of population and economy for that city in China. Most Chinese cities are not formed following administrative boundaries, but are a result of thousands of years of history and cultural development. The cities in China tend to be defined more likely by human and economic clusters than by arbitrary boundaries. Therefore we believe that for Chinese cities it is appropriate to use the definition of a city supplied for administrative purposes.

[8] Taking the great four ancient capitals for example, Beijing, Nanjing, Xi'an and Luoyang each has over 2000 years of history of human settlements.

Chinese city level annual population data from 1984 to 2009 are provided by the '*China Urban Statistical Yearbooks' (National Bureau of Statistics of China, 1985-2010).* Data in the yearbook series extend back to 1984 and the city definitions obey the 1984 criteria[9], adjusted to year 2000 statistical changes. Hong Kong, Macao and Taiwan are excluded. Note that there are four urban population variables reported for each city.

- ***Type A:*** Total Population of the Entire City (*'Diqu'*[10]). A term used by the NBSC[11] since 1985 to refer to the total population of the whole city. More clearly, the so-called 'entire city' is not really a city; it is a city district consisting of one administering central city and several counties under the administration of the central city. Therefore, Chongqing with a total population of 32,756,100 inhabitants is China's largest city in 2009, which consists of one central city-proper area and 19 counties[12].

---

[9] From the 1983 Chinese urban system launched the policy of 'city governing the surrounding towns', the central government assigned some towns to the adjacent city in order to take the advantage of the city to help the development of the towns.

[10] From 2002, 'China Urban Statistic Yearbook' has employed the new term 'Quanshi' and 'Shixiaqu' to represent 'Diqu' and 'Shiqu' respectively, the meanings does not change. In this paper we still use 'Diqu' and 'Shiqu', because the meaning of these two terms are more explicit.

[11] NBSC: National Bureau of Statistics of China.

[12] Counties in China consist of small villages which mainly conduct agriculture activities or basic small handcraft industries and do not have the same infrastructures and amenities as cites. They are not cities but their population are included in Type A data in 'China Urban Statistical Yearbook', hence we do not use Type A data in our paper.

- ***Type B:*** Total Population of City Proper (*'Shiqu'*). This is the data we use in this paper, because it reports the real urban population (resident) who actually live in the urban area and benefit from the urban infrastructure and amenities[13]. So Chongqing's urban area population is reported as 15,427,700 inhabitants, nearly half of the ***Type A*** data. It is noteworthy that besides the urban population, these type of data include the population in adjacent suburbs; see Figure 2.1 which indicates the models of administrative structure of Chinese cities proposed by Ma & Cui (1987).[14] Model (a) shows the 'city proper'- the basic administrative structure of the core of cities where urban populations are concentrated in the urban core and surrounded by adjacent suburban areas. Model (b) & (c) illustrates the typical entire city districts, as ***Type A*** data. An entire city consists of one city proper (includes the adjacent suburbs) and one or several administered counties.

- ***Type C:*** Non-agriculture Population of the Entire City. A component of ***Type A*** data, which accounts for the population within an Entire City district holding non-agriculture *'Hukou'*.

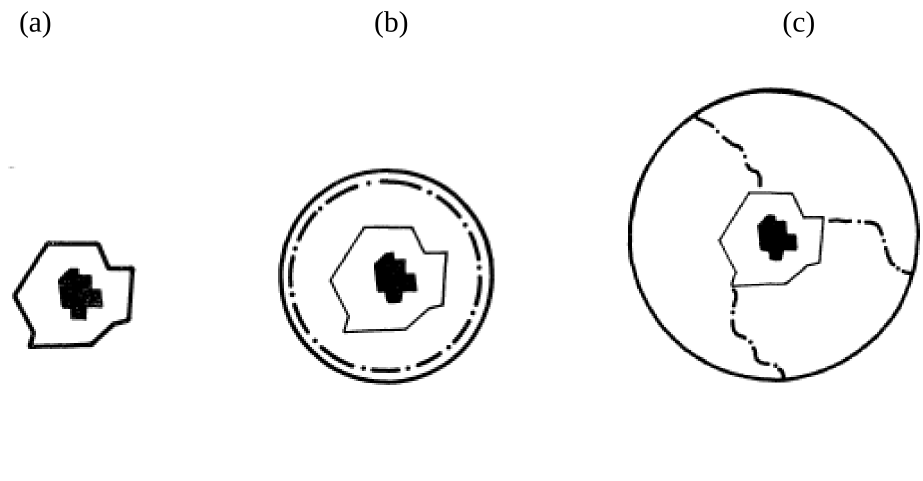
- ***Type D:*** Non-agriculture Population of City Proper. A component of Type ***B*** data, which accounts for the population within a City Proper holding non-

[13] This may include residence with a 'hukou' affiliated to that city or residence temporarily living and working in that city (residence with a 'hukou' affiliated to some other city).

[14] Though it is the model in 1984, until now there is no major change of Chinese urban system since 1984 when the urban system was established.

agriculture *'Hukou'*. In China, these kind of data refer to the urban district only in Figure 2.1 mode (a), it would exclude the outer suburban area surrounded. Because in China, unlike developed countries, the residents living in the adjacent suburbs are not relatively rich workers who dwell in the suburbs and do not all work in the urban core, but may actually work in the agricultural sector. Therefore, some studies use this kind of data to conduct the studies for Chinese urban distribution, like (Song and Zhang, 2002; Xu and Zhu, 2009; Wang and Zhu, 2012).

With respect to county-level cities, they do not administer any county and plainly take the structural as 'city proper' shown in Figure 2.1 model (a). Therefore in *'China Urban Statistical Yearbook',* the population for county-level cities is only reported in Type B data.



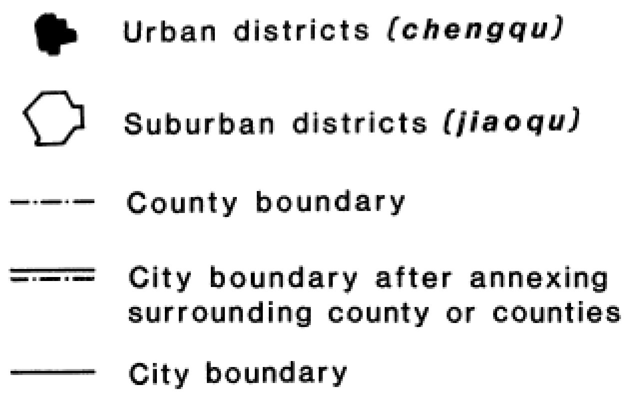


Figure 2.1: Generalized model of administrative structures of Chinese cities, since 1984. (Source: Ma & Cui, 1987)

We choose Type B data in this paper since we believe this represents the real urban population data. Because on the one hand, governing the surrounding towns is not the main function of a city; all the functions of a city are reflected in the 'city proper' data ('*Shiqu'*). On the other hand, the amount of towns that are governed by a city is administratively assigned and may vary over time (e.g. sometimes a town may separate from a city and upgrade into an independent city), hence the second type of data- 'city proper'- are relatively stable and convenient for the comparison either with the city itself or foreign cities. With regard to the Type C & Type D data- 'non-agriculture population', after the 'Economic Reform of 1979" and concomitant rural reform, the effect of the *'hukou' ID system* is fading out, there is a considerable amount of temporal migrants (who are free from their farm work) attracted to the urban district ('city proper') to seek urban job opportunities. Hence, it would generate an obvious underestimate for an urban size if the accounting is simply based on non-agricultural population. Appendix Table A2.1 lists these four types of data as in '*China Urban Statistical Yearbooks'* of 655 Chinese cities in 2008.

The city size annual data (population) before 1984 (1879 to 1983) comes from the most comprehensive dataset of world urban populations, by Jan Lahmeyer (*http://www.populstat.info/)*. These data are also consistent with Type B data from the *'China Urban Statistical Yearbook'*. However, the historical population data in some years is not complete due to certain external conditions (World War II) and census technologies during that period.

2.2 CHINESE URBAN SYSTEM

The Chinese urban system is quite complicated. In China (Pepole's Republic of China)[15], in terms of the hierarchy of administrative levels from highest to the lowest, there are three different administrative levels of cities in the urban system: (1) provincial-level city, which is under the direct administration of central government. (2) prefecture-level city[16], which is under the administration of relative province

[15] These administrative levels of cities all refer to the urban system that developed after the PRC was established in 1949. For the period before 1949, there is no systematic and stable urban system due to the complicated political situation.

[16] Actually, there is a 'sub-provincial level city' between provincial-level city and prefecture level city since 1994, which was planned to be provincial-level cities. However, this kind of cities are divided due to administrative or political purpose and the characteristics of these cities are not different from

(including the province capital). They are the main body of cities in China, 268 prefecture-level cities in 2009; (3) county-level cites, normally under the administration of adjacent prefecture-level cities but some may directly under the administration of the relative province. There are 368 county-level cities in 2009. This paper involves all of these three level of cities.

Specifically, there are 4 provincial-level (directly controlled municipality) cities in China by 2009, Beijing, Tianjin, Shanghai and Chongqing, each taking the administrative role as a province. Taking Shanghai for example, in terms of administrative level, it is a provincial-level city which means Shanghai is administered directly by central government and has the status of a province. However, the provincial-level cities still have the same administrative structure like other prefecture-level cities, within a city's jurisdiction, there are several urban districts which consist of the urban area/ city-proper (indicated as *'Diqu'* in *'China Urban Statistical Yearbook'*) and a few counties. For example, Shanghai has 16 urban districts [17] and one county (Chongming County) under its jurisdiction, hence the two levels of population data reported in *'China Urban Statistic Yearbook'*-*'Diqu'* and

---

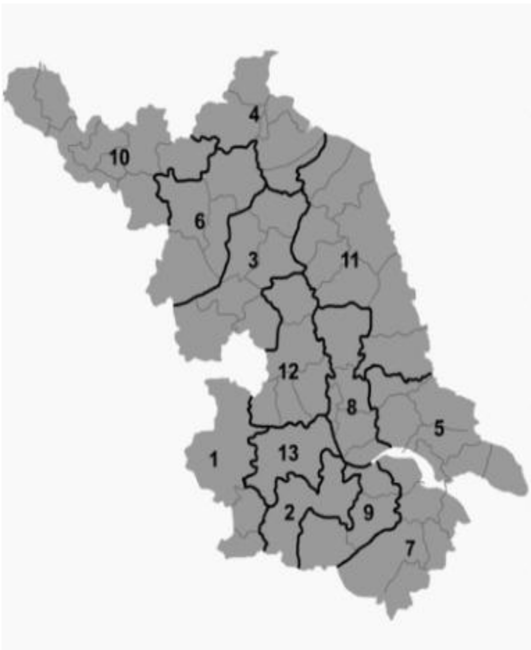
prefecture level cities, therefore we divide this these sub-provincial-level cities into prefecture-level cities in this paper. Until 2009, there are 15 sub-provincial level cities: Shenyang, Dalian, Changchun, Harbin, in North-East region; Nanjing, Hangzhou, Ningbo, Xiamen, Jinan, Qingdao, Guangzhou, Shenzhen in East region; Wuhan, Chengdu, Xi'an in Mid-land region.

[17] Shanghai's 16 urban districts are: Huangpu, Xuhui, Changning, Jingan, Putuo, Zhabei, Hongkou, Yangpu, Minxing, Baoshan, Jiading, Pudongxinqu, Jinshan, Songjiang, Qingpu, Fengxian districts, which consitute the urban area of Shanghai.

'Shiqu', means the total population of the 16 urban districts plus one county and the population of these 16 urban districts only.

With respect to prefecture-level cities, they are under the jurisdiction of the relevant province (including the province capital) and constitute the main body of the Chinese Urban system. Taking Jiangsu province for example (the neighbor of Shanghai, Figure 2.2), in 2009 there are 13 prefecture-level cities in Jiangsu Province (including one province capital, Nanjing) and 26 county-level cities. Figure 2.3 shows the boundaries of these 13 prefecture-level cities/ districts with the relative city population on the right (the whole city district population- *'Diqu'*). Moreover, Figure 2.4 depicts the city proper of these 13 prefecture-level cities. Each shaded area is the 'city proper' area for the corresponding city district, within a city district, besides the shaded area are the counties that the city governs or where a county-level city located (which I do not show in this figure).

Figure 2.2: The location of Jiangsu Province in China, 2009. (from itourschina.com)



#	Name	Chinese name	Population 2009 (1000 inhabitants)
1	Nanjing*	南京	6,297.70
2	Changzhou	常州	3,598.20
3	Huai'an	淮安	5,341.60
4	Lianyungang	连云港	4,906.40
5	Nantong	南通	7,626.60
6	Suqian	宿迁	5,406
7	Suzhou	苏州	6,332.90
8	Taizhou	泰州	5,039.80
9	Wuxi	无锡	4,656.50
10	Xuzhou	徐州	9,576.10
11	Yancheng	盐城	8,123.70
12	Yangzhou	扬州	4,588
13	Zhenjiang	镇江	2,698.80
* means the province capital			
population refers to end of year total city related area population- *'Diqu'*			

Figure 2.3 Jiangsu Province is divided into 13 prefecture-level cities/ divisions.

Figure 2.4 Administrative system of Jiangsu Province, 1984. (Source: Ma & Cui, 1987)

County-level cities formed from two sources, one is the update of the towns and the other is converted from counties or districts. After 1980s, with the economic development, the urbanization process is rapid and most of the county-level cities come from the latter source. However, in 1994 the central government forbade the update from town, counties or districts to avoid the acceleration of diminishing farmland. After a decade of development, the forbidding of updating to cities cannot mitigate the diminishing of farmland but there emerged a lot economic developed counties and large urban population counties. Therefore, in 2005, the central government relaxed the restriction on upgrading from counties to cities and issued a criterion of updating from a county to a city, including the standards for GDP and the composition of GDP; the population size where the local government locates and the composition of non-agriculture inhabitant and agriculture inhabitant; the local amenities including the popularity rate of tap water, green area per capita, sewage disposal ratio.

2.2.1 Four Economic Regions

The National Bureau divided China into four economic regions on 13th June 2011: East, Midland, West and North-East[18]. We group data according to these definitions in later regressions. See Figure 2.5 for the boundary of four economic regions. From left to right they are West, Midland, East, and North-East economic regions. Following 'Economic Reform' from 1979 most of the resources were targetted to the East and then from 2000 there was a national policy of 'Go West', followed by 'The Manchurian Mandate' in 2004 and 'Rise of the Midland' in 2006[19].

[18] The east region includes Beijing, Tianjin, Shanghai and seven provinces; the Midland region includes six provinces; the West includes Chongqing and 11 provinces; and the North-East includes 3 provinces. See Figure 2.5 Before 2011, there are three economic regions: East, Mid-land and West.

[19] More information can be find in http://international.fhwa.dot.gov/pubs/pl08020/fmic_08_02.cfm

Figure 2.5 Four economic regions of China

Source: National Bureau of Statistics of China. Accorindg to the 13th June 2011 documents of the National Bureau, which divide China into four economic regions.

From left to right they are West, Midland, East, North-East economic regions and the relative policy is at the beginning of the 'Economic Reform' most of the resources will support the East region to develop first, then from 2000 there is a national policy of 'Western Development', then 'The Manchurian Mandate' in 2004 and 'Rise of the Midland' in 2006.

2.2.2 Overview of Chinese Urban Development

During the period of our study 1879-2009, there are two significant different stages of urbanization development under certain political circumstances. The year 1949 is the People's Republic of China's foundation year; therefore there will be different stories before and after 1949. For the years before 1949, firstly, from 1879 to 1912 it was the end of the Qing Dynasty and from 1912 to 1949 China was ruled by the nationalist party (Kuomintang, KMT). Although both of these two periods experienced wars and upheaval, from 1879 to 1926 both the number of cities and urban population were relatively stable (see Figure 2.6), because the urban development were not the primary concern for that period, and cities in that period were affected the least from the wars because there were many foreign concessions in the cities. Then from the year 1926, the number of cities more than doubled in the next decade from 31 in 1926 to 82 in 1936 and increased slightly until 1948 when there were 88 cities in China. From 1926 the rule of the nationalist party matured and the population grew steadily between 1926 and 1948.

Under the rule of the communist Party from 1949, urban development has increased dramatically. The number of cities increased from 132 in the beginning year 1949 to 653 in 2009; and the urbanization rate rose from 7.3% in 1949 to 45.68% in 2008. Generally speaking, Chinese urban development has gone through 5 stages, which should be borne in mind when we analyse our results in section 4.

(1) Starting of urbanization (1949-1957)

As shown in Table 2.1 below, there were only 132 cities in 1949 and 39.49 million urban residents (7.3% of total population), after the *'first-five year plan'* a number of new mining and heavy-industry cities emerged and the old cities and medium-sized cities were improved. At the end of year 1957, there were 176 cities, an increase of 33% compared to 1949, and urban population had risen to 70.78 million, an increase of 79.2% compared to 1949.

(2) Huge fluctuation stage (1958-1965)

During the *'second-five year plan'*, urban development showed huge fluctuation from fast expansion to contraction because of the fast expansion and contraction of the national economy. After the 3 years of the *'Great Leap Forward'*, the number of cities increased from 176 (1957) to 208 (1961), an increase of 18.2%; urban population rose from 70.78 million to 101.32 million, an increase of 43.2%; the urbanization rate expanded from 10.9% to 15.4%. The *'Adjustment of National Economy'* started from 1962 and restricted the number of cities, by the year of 1965; there were 168 cities, 40 cities reduced from 1961 (decreased by 20%). The main reason for this contraction was the policy that required some newly designed cities to return to counties. A large number of construction projects were stopped or delayed and 25 million workers in cities were forced to go back to counties and villages. The urban population reduced to 88.57 million in 1965 from 101.32 million in 1961 (decreased by 12.6%); urbanization rate shrank from 15.4% to 12.2%. These are purely administration changes.

(3) Stagnant urban development (1966-1978)

The *'Cultural Revolution'* started from 1966 and lasted at least 10 years, during this period the national economy remained stagnant, as did urban development. From 1966 to 1978, the number of cities only increased by 26, 2 cities emerged every year on average. However, data show that the urban population increased by 94.69% from

88.58 million in 1965 to 172.45 million during this period, indicating that urban development in this period was focused on the expansion of existing cities.

(4) Rapid urban development (1979-1991)

Since the ‘Chinese Economic Reform’ launched in 1979 with the implementation of a series of reform policies on domestic economy and openness to the world, the urban development and national economy experienced rapid growth. In the 1980s, economic reforms spread out in succession. Especially in 1990s, the policies that favor to small & medium-sized cities and the rise of town-ship enterprises both accelerate the rapid growth of urban development. From 1979 to 1991, during these 12 years the number of cities increased by 286, 15 cities emerged ever year on average. By the end of 1991, the urban population had expanded to 312.03 million, an increase of 80.9% compared to 1978; the urbanization rate was 26.94%, a 9% increase from 1978.

(5) Relatively Stable urban development (1992-2009)

From Figure 2.6 below, we can see that the number of cities was relatively stable during this period and urban population increased steadily. Specifically, with a relaxation of migration between cities and especially between rural and urban area, the migration increased consequently. At the end of 2009, compared to 1991 the number of cities reached 655, 176 more cities emerged, 11 new cities every year on average; urban population increased by 100.44% and the urbanization rate increased by 20%.

Table 2.1: Description of 5 stages of Chinese urban development

year	no. of cities	Growth rate comparing to previous stage (%)	urban population (million)	Growth rate comparing to previous stage (%)	Urbanization rate (%)
1949	132		39.49		7.3%
1957	176	33.34%	70.7727	79.22%	10.9%
1961	208	18.18%	101.3247	43.17%	15.4%
1965	168	-19.23%	88.5762	-12.58%	12.2%
1978	193	14.88%	172.45	94.69%	17.92
1991	479	147.67%	312.03	80.94%	26.94%
2009	655	37.03%	625.4391	100.44%	46.87%

Source: *'China Urban Statistical Yearbook-2009'* Urbanization rate= urban population/ total population

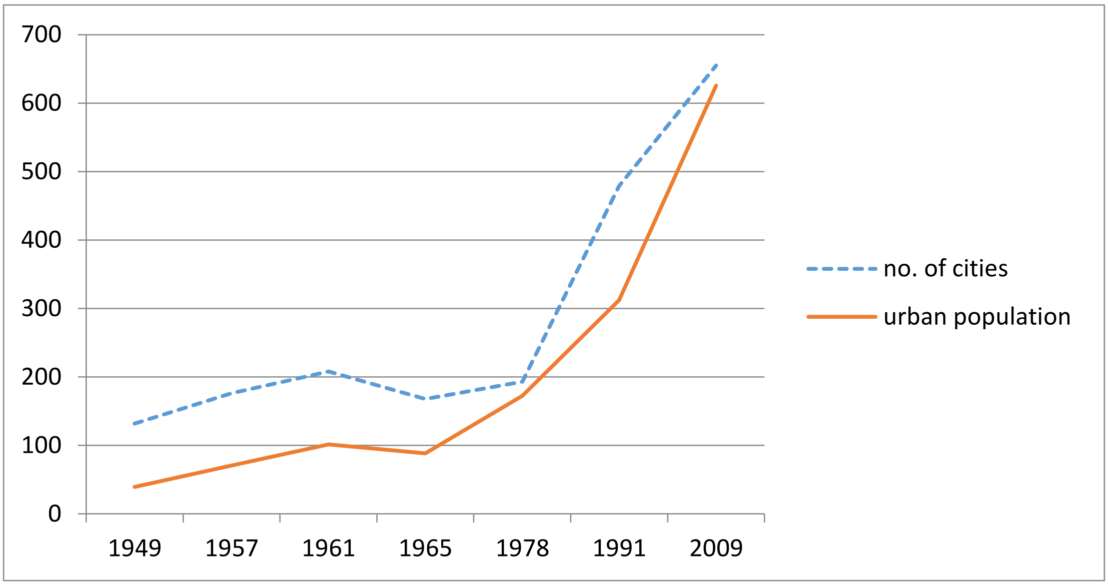


Figure 2.6 Evolution of urban growth over 5 stages mentioned above.

To conclude, Table 2.2, Figure 2.7[20] and Figure 2.8 below provide a brief statistical summary of the population data used in this paper- Chinese annual end-of-year urban population from 1879 to 2009. The total population of China increase from 366.99 million in 1879 to 1,334.5 million in 2009 (which now accounts for nearly one fifth of

[20] Figure 2.7 is the long run compared to Figure 2.6. The number of cities seems increased rapidly in short run Figure 2.6, while the increase is relatively small if one looks at the longer run in Figure 2.7.

the world population), a huge increase of 967.51 million population to the planet; while the urban population grow from 11.4 million to 645.12 million, an increase of 633.72 million urban population over last century accounting for 65.5% of the total population growth..

The data of urban population between 1959 and 1982 is not available; however, it will not affect our results too much. During this period China was experiencing urban development stages (2) & (3) mentioned above: the huge fluctuation (1958-1965) and stagnant (1966-1978). After these periods, the ‘Economic Reform’ launched in 1979 but had effect a few years later in the early 1980s.

2.2.3 Review of three policies in China

There are three policies might influencing the city size distribution in China: ‘Economic Reform’, ‘One Child Policy’ and ‘Hukou system’, which also discussed in the following chapters in terms of how these unique policies in China affect the validity of Zipf’s law and Gibrat’s law. We state the purpose and timeframe of the policies first and then we generally discuss the impact of these policies on Chinese city size distribution (more specific discussions can be found in the following chapters associated with the results).

- ***'Economic Reform'***

'Economic Reform' was launched in 1979 and aimed at promoting the economic development in terms of changing from command economy to market economy, open up to trade, loosened the restrictions on intercity migration. From 1979 until 2010, unprecedented growth occurred, with the economy increasing average by 9.5% a year. The private sector grew remarkably, accounting for as much as 70% of China gross domestic product by 2005 (China Statistical yearbook).

The reform was carried out in two stages. The first stage was supporting prior for some specific regions or cities to develop first, for instance, constructing some 'special economic zones'. The second stage was the expansion of these developed regions to the overall country. These two stages are also reflected in the Zipf's exponent.

- ***'One Child Policy'***

Almost coincidentally with 'Economic Reform', 'One Child Policy' was introduced in 1978. It is a population control policy of the People's Republic of China. The aim of the policy is controlling for the rapid growth of population, as the population grew from around 540 million in 1949 to 940 million in 1976 (China Statistic Yearbook). After two decades, as some social problems increasingly serious (ageing population, etc.), the majority of provinces and cities permit two parents who were 'only children' themselves to have two children and the policy was officially relaxed in 2010. In 2013, this rule was relaxed even further: couples in which one parent is an only child are allowed to have a second child (The Economist 2015).

- ***'Hukou system'***

'Hukou system' is a record of household registration required by law in People's Republic of China. Individuals were broadly categorised as a "rural" or "urban" worker. A worker seeking to move from the country to urban areas to take up non-agricultural work would have to apply through the relevant bureaucracies. The number of workers allowed to make such moves was tightly controlled.

In 1958, the Chinese government officially promulgated the family register system to control the movement of people between urban and rural areas. The original aim is that with large rural population of poor farm workers, 'Hukou' limited mass migration from the land to the cities to ensure some structural stability.

However, after the 'Economic Reform 1979', in practice the system has largely broken down. Because 'Economic Reform' created pressures to encourage migration from the interior to the coast (the coast regions were the first to develop according to the policy of ‘Economic Reform’). It also provided incentives for officials not to enforce regulations on migration to gain economic development.

- ***Possible linkage between policies and city size distribution***

‘Economic Reform’ carried out from 1979 until nowadays. Basically, our results in following chapter show that ‘Economic Reform 1979’ may help producing more equal cities, i.e. after ‘Economic Reform’ city size distribution tends to be more evenly, as the Zipf’s exponent generally increases after the reform. The possible explanation could be that the ‘Economic Reform’ affects city size distribution through the increasing level of economic development. The ‘Economic Reform’ greatly promotes the national economic growth and improves the income level substantially. The decreasing disparity of income level may lead to the decreasing population differences among cities. This is also consistent with Anderson and Ge (2005).

‘One Child Policy’ starts from 1978 and officially relaxed in 2010. Although it changed fundamentally the natural growth rate of population, the policy is enforced at the provincial level in urban areas, thus the shock are the same for all the cities in our sample (only lower the variance of city growth), i.e. cities still grow under the identical growth process where Gibrat’s law may hold and Zipf’s law may emerge in the steady state. Therefore, this policy may not affect city size distribution.

‘Hukou’ system strictly restricted the migration from rural to city during the period 1958 to the end of 1980s, but not the same for the intra-city migration. Thus, in terms of the influence of city size distribution, this policy is the same as ‘One Child Policy’.

To conclude, in fact as all these three policies are the shocks for all cities, technically, all of them do not affect the validity of Zipf’s law or Gibrat’s law. Because if the shocks are applied to all cities, then cities still grow under identical growth process, where Gibrat’s law may hold and then Zipf’s law may also hold in the steady state. However, ‘Economic Reform’ may indeed promote more evenly distributed cities as it improve the level of economic development and the level of income for all cities. This may lead to the decrease of differences between city sizes, as cities are more identical for people to live for a better live (better income[21]).

Although these policies may not affect the validity of Zipf’s law or Gibrat’s law, we still discuss these policies because we are investigating city size distribution within China, while these policies are unique in China compared with studies focusing on other countries.

[21] We do not consider the environment quality as the determinants of location choice of residence at this stage to address the current question more clearly.

Examining the link between these policies to the validity of Zipf's law and Gibrat's law is interesting, but also challenging as there are still no exact literature investigating the precise relationship between policies and the validity of Zipf's law and Gibrat's law. In this thesis, we simply suggest how these policies may affect these empirical regularities.

We mainly focus on the empirical status of city size distribution of Chinese cities, i.e. whether the Chinese city size distribution reaches a steady state (Zipf's law), if not, how far is Chinese city size distribution below or above the steady state (Zipf's law). This is important as the city size distribution is attracting increasingly attention in regional studies, i.e., if population spreads too much and the size of single city will be too small, then it is hard to take the advantage of scale economy, and hard to construct close connections among cities which may waste the infrastructure construction; in contrast if population concentrates only in a few large cities, this will lead diseconomies of scale and congestion, environment degradation, housing shortage, employment difficulties, etc.

However, as we simply imply the possible mechanism of how these policies may influence city size distribution, we have to acknowledging that one does not have enough evidence to claim that these effects are indeed in place. More work on the precise relationship between the policy and the validity of the Zipf's law and Gibrat's law could be done in the future.

Table 2.2: Data description

Year	No. of cities	Prefecture level cities and above	county-level cities	Min. city size (1000 persons)	Max. city size (1000 persons)	Mean city size (1000 persons)	Standard deviation	urban population (million)	totoal population (million)	Urbanization rate
1879	29	28	1	20	1,648.80	393.12	363.01	11.4004	366.99	3.11%
1911	37	35	2	20	900	263.85	268.61	9.7625	427.66	2.28%
1918	30	28	2	15.5	1,444.00	267.22	355.76	8.0167	461.77	1.74%
1926	31	30	1	20	1,583.90	375.40	413.58	11.6375	482.13	2.41%
1936	82	80	2	32	3,490.00	290.05	454.59	23.7842	507.86	4.68%
1948	88	86	2	7	4,423.00	356.51	566.47	31.373	547.80	5.73%
1953	166	128	38	30	6,204.40	332.40	656.56	49.7993	587.96	8.47%
1958	175	114	61	43	6,977.00	521.05	847.22	65.1313	654.16	9.96%
1983	281	140	141	73.7	7,551.20	706.53	907.77	222.74	1,023.30	21.62%
1984	295	146	149	6.8	6,881.30	648.97	779.97	240.17	1,043.57	23.01%
1985	324	165	159	7.2	6,983.00	655.17	761.15	250.94	1,058.51	23.71%
1986	353	169	184	7.7	7,101.60	660.42	777.31	263.66	1,075.07	24.52%
1987	381	173	208	8.1	7,217.70	686.67	755.17	276.74	1,093.00	25.32%
1988	434	186	248	8.3	7,326.50	687.14	726.60	286.61	1,110.26	25.81%
1989	450	188	262	9	7,777.90	705.22	734.56	295.40	1,127.04	26.21%
1990	467	188	279	9.6	7,834.80	718.26	732.15	301.95	1,143.33	26.41%

1991	479	190	289	9.9	7,861.80	723.89	730.44	312.03	1,158.23	26.94%
1992	517	194	323	10.3	7,927.50	735.62	714.31	321.75	1,171.71	27.46%
1993	570	199	371	10.8	9,480.10	755.64	731.98	331.73	1,185.17	27.99%
1994	622	209	413	11.2	9,530.40	768.96	716.10	341.69	1,198.50	28.51%
1995	640	213	427	11.5	9,566.60	781.05	737.88	351.74	1,211.21	29.04%
1996	666	221	445	11.8	9,610.20	775.57	735.28	373.04	1,223.89	30.48%
1997	664	222	442	–	–	–	–	394.49	1,236.26	31.91%
1998	664	227	437	–	–	–	–	416.08	1,247.61	33.35%
1999	690	263	427	16	11,272.20	791.91	886.55	437.48	1,257.86	34.78%
2000	663	263	400	16	11,368.20	839.33	926.61	459.06	1,267.43	36.22%
2001	653	260	393	18	12,624.10	847.21	963.81	480.64	1,276.27	37.66%
2002	660	279	381	19	12,702.20	878.15	1019.20	502.12	1,284.53	39.09%
2003	660	286	374	22	12,782.30	891.79	1041.03	523.76	1,292.27	40.53%
2004	661	287	374	31	12,891.30	904.67	1057.13	542.83	1,299.88	41.76%
2005	661	287	374	47	12,901.40	920.69	1090.91	562.12	1,307.56	42.99%
2006	656	287	369	50	15,109.90	929.60	1138.04	582.88	1,314.48	44.34%
2007	655	287	368	47	15,260.20	940.08	1156.32	606.33	1,321.29	45.89%
2008	655	287	368	47	15,345.00	947.71	1166.46	624.03	1,328.02	46.99%
2009	655	287	368	48	15,427.70	957.02	1180.05	645.12	1,334.50	48.34%

Source: data before 1983 (inclusive) is from *Jan J. Lahmeyer http://www.populstat.info/* ; Data after 1984 (inclusive) is from *'China Urban Statistical Yearbook', National Bureau of Statistics of China, 1985-2010; and '60 Years of PRC', NBSC 2010; 'Almanac of China's Population 2011', Institution of Population and Labour Economics, CASS.*

For data from 1959 to 1982, only some important cities data are available. We will not use it for the Zipf's exponent, but only for the panel data in testing for Gibrat's law. Besides, for this period China urban system was under huge fluctuation and stagnant stage (see section 2.3.2), it will not affect the results much if we do not taking into this period for Zipf's law.

For data of 1997 and 1998, only the data for prefecture-level cities are available in all kinds of yearbooks, still we will use it in the panel data testing for Gibrat's law, not in testing of Zipf's law.

For data from 1999 to 2004 (inclusive), we found that for these 6 years, the county-level city population data is not consistent with the other years in *'China Urban Statistical Yearbook'*. As for these 6 years they have both the specific population data of 'city proper' and the whole city area data available, while for other years the city proper data is not available, for 'city proper' column they use the same data of whole county-level city area data instead. Therefore to make sure the consistency of data, we use the whole city area data for county-level cities in 1999-2004.

The two datasets could merge well, there is no big gaps between these two data sources, see Appendix FigureA2.1.

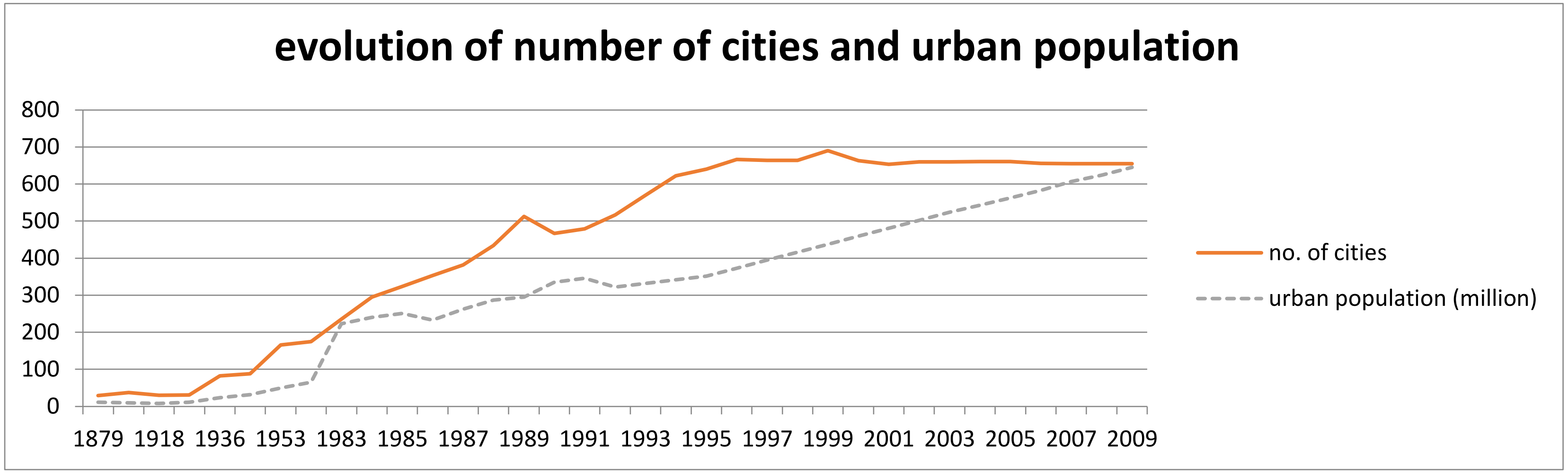


Figure 2.7 Growth in urban population and number of cities, 1879-2009

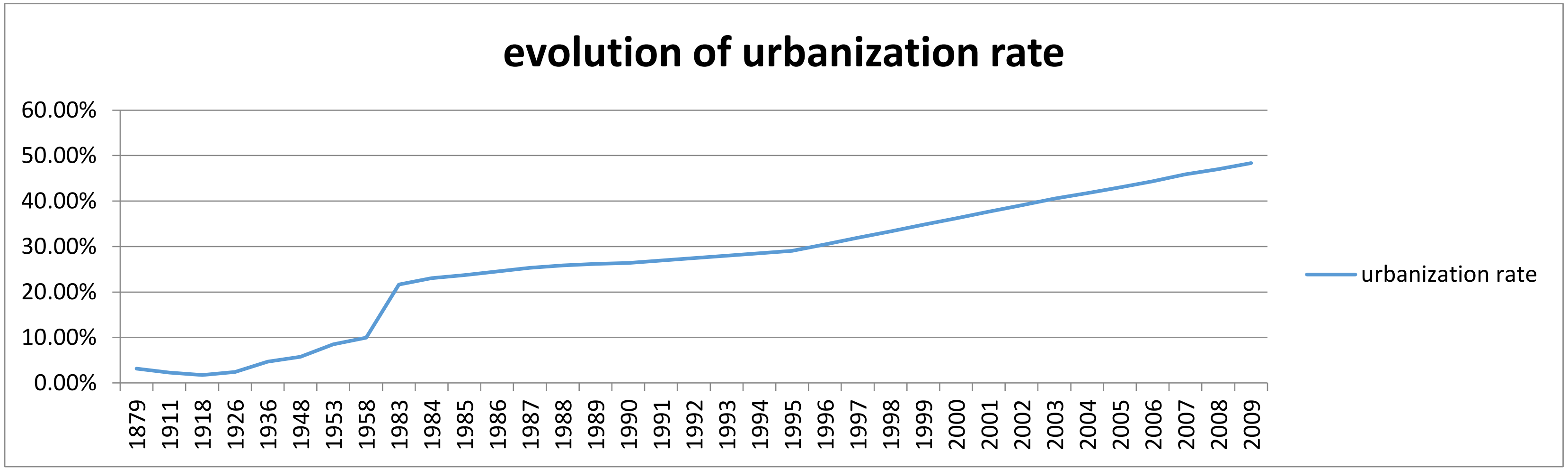


Figure 2.8 Evolution of urbanization rate, 1879-2009

CHAPTER 3
CHINESE CITY SIZE DISTRIBUTION: TESTING FOR ZIPF'S LAW

3.1 INTRODUCTION

This paper revisits the well-known empirical regularity- Zipf's law and applies it to Chinese cities, i.e. whether Chinese city size distribution could be approximated by a power law [22] distribution (Zipf's law). According to the recent report of the 'Sixth Census of China 2010' released in April 2011, nearly 0.666 billion of Chinese people live in cities (including county-level cities), which means nearly half of the Chinese population is living in urban areas. Despite such rapid growth, new construction and urban sprawl continues. Against a background of rapid urbanisation it is important for acdemics and poliymakers to understand how cities develop and how their distribution changes over time. A knowledge of city size distribution will help policymakers make informed decisions on, for example, urban infrastructure, education and health investment or environment management. Our study of city size distribution can also contribute to the acdemic study of the stylized empirical regularity- 'Zipf's law'.

Specifically, Zipf's law considers city size (size refers to the population) and city rank (by population), and shows that the rank is perfectly inversely correlated to its size, at

[22] A power law is a mathematical relationship between two quantities. When the frequency of an event varies as a power of some attribute of that event (e.g. its size), the frequency is said to follow a power law. There is evidence that the distributions of a wide variety of physical, biological and man-made phenomena follow a power law, including the size of earthquakes, craters on the moon and of solar flares, the foraging pattern of various species, the size activity patterns of neuronal populations, the frequencies of words in most languages, frequencies of family names and many other quantities.

least for large cities. More detailed, Zipf's law argues that there is a relationship like city rank $\mathrm{r} = \mathrm{c}p_{(r)}^{-1}$, where $p_{(r)}$ is the population of rank r. Then after simple derivaration, $p_{(r)} = (c/r)^{1}$, thus $\frac{\boldsymbol{p_{(1)}}}{\boldsymbol{p_{(2)}}} = \frac{\boldsymbol{c/1}}{\boldsymbol{c/2}} = \mathbf{2}$ and therefore $\frac{\boldsymbol{p_{(1)}}}{\boldsymbol{p_{(r)}}} = \frac{\boldsymbol{c/1}}{\boldsymbol{c/r}} = \frac{\boldsymbol{r}}{\mathbf{1}}$. This implies that the first ranked city has roughly twice the population of the 2nd ranked city, has roughtly three times the population of the 3rd ranked city, etc. This suprising Zipf's distribution for cities has been proved empirically across many countries and over time.

Strictly speaking, Zipf's law is not a desirable or optimal city size distribution; it more seems like a long run stable status of distribution of city sizes. The importance of this law is that, given very strong empirical support, it constitutes a minimum criterion of admissibility for any model of local growth, or any model of cities (Gabaix, 1999). In this thesis, we test for Zipf's law not only to confirm the validity of Zipf's law, we also attempt to reveal the status of distribution of populations in China. Because with the rapid development of economy and urbanisation in China in recent decades, whether we should lead the increasing urban population to concentrate in large cities or it is better to spread out in mediate-sized and small cities? Investigating the city size distribution is important. If population spreads too much and the size of single city will be too small, then it is hard to take the advantage of scale economy, and hard to construct close connections among cities which may waste the infrastructure construction; in contrast if population concentrates only in a few large cities, this will lead diseconomies of scale and congestion, environment degradation, housing shortage, employment difficulties, etc.

We not only test for the validity of Zipf's law in China, but also use the Zipf's exponent as a benchmark of whether the city size has been evenly distributed. If Zipf's exponent greater than 1 or increasing, then city sizes are more evenly distributed and getting more and more even. This also reflects the expansion of mediate-sized and small cities. In addition to Zipf's exponent, we also mention urban primacy (the proportion of largest cities on total urban population) to investigate the urban development. The primacy is expected to be negatively correlated with Zipf's exponent, and the results confirm this.

In addition, Zipf's law and Gibrat's law (in the next chapter) offer a strong benchmark against which to measure theories of urban evolution and to organize an up to date look at the literature. The robustness of Zipf's law has also served to attract attention to the need for microfoundations (Gabaix and Ioannides, 2003).

The contribution of this paper is twofold. Firstly we construct a unique dataset that, for the first time, allows us to investigate the evolution of Chinese city size distribution and the growth pattern for over a century (1879-2009). Most of the previous empirical literature testing for Zipf's law on Chinese cities uses only a few years, and the earliest year used in the literature is 1949 (Anderson and Ge, 2005). According to the previous literature the results of testing for the city size distribution and urban growth pattern tend to be highly sensitive to the year involved. That is why some studies on Chinese city size distribution support Zipf's law but some do not (Song and Zhang, 2002; Gan, Li and Song, 2006; Soo, 2014; Schaffar *et al*., 2010 support Zipf's law while Anderson and Ge, 2005; Xu & Zhu, 2009; Peng, 2010; Ye and Xie, 2012 do not). Our results close the debate to some extent as showing that Chinese city size distribution consistently becomes more and more even from 1879 to 2009, with some evidence of Zipf's law in the late 1980s and early 1990s. We find that Chinese city size distribution becomes more even (equal) than Zipf's prediction as time passes, i.e. the disparity between large and small cities is smaller and smaller than Zipf's law predicts over time.

Secondly, for the first time for Chinese cities, the whole sample is divided into prefecture-level cities and county-level cities according to cities' different administrative functions, and then divide cities into four economic regions (East, Central, West and North-East cities) as the unbalanced economic development in different regions. In addition, the most unique part of our paper is that we also divide our sample into four historical subgroups according to city age: 29 historical cities from 1879 to 2009, 82 historical cities from 1936 to 2009, 125 historical cities from 1958 to 2009 and 294 cities from 1984 to 2009, which can highlight the old and consistent cities. This kind of taxonomy enables our study of Zipf's law to be more suitable for Chinese cities' unique conditions. We are trying to see whether the size distribution predicted by Zipf's law holds for each sample. (1) As expected, Zipf's law holds for prefecture-level cities as prefecture-level cities are the cities with policy priority of development and have long been the human agglomeration and economic centre. This is consistent with the empirical literature that finds Zipf's law holds for cities at least large cities, because the definition of prefecture-level cities in China is close to the international definition of a city used in the previous literature. In contrast, county-level cities are distributed much less evenly than Zipf's law predicts, but they become more and more even over time (Pareto exponent increases over time) and stay relatively stable after 2000 with a Pareto exponent value of -0.8, which shows county-level city size distribution still less equal than Zipf's distribution, i.e. the disparity between large and small cities is relatively big. (2) Zipf's law is found in the East region cities' distribution. Other regions do not support Zipf's law. (3) Zipf's law also seems to hold in the balanced data of 82 historical cities (1936) sample with Pareto exponent -0.977 (result also shows that this is not statistically significantly different from -1).

This paper is organized as follows. Section 2 presents the emergence of Zipf's law, together with a review of the empirical debate on city size distribution. Section 3 outlines the methodology used for testing for Zipf's law. Section 4 shows the results. The final section concludes.

3.2 LITERATURE REVIEW

City size distribution has been subject to numerous empirical studies in various fields like urban economics, statistical physics and urban geography. Zipf's law, a surprising empirical regularity, describing the shape of city size distribution as a special evenly distributed pattern, i.e. the largest (rank 1) city's population is twice the second (rank 2) largest city's population and three times the third (rank 3) largest city's population etc., attracts lots of attention in urban economics. In recent years there have been a large number of empirical works related to Zipf's law, and a debate has ensued about whether Zip's law (indicates Pareto distribution) can really approximate the city size distribution. Because Pareto distribution (indicated by Zipf's law) is found in various countries, however, recently log-normal distribution (indicated by Gibrat's law) is found when one includes all the cities or places (actually the definition of a city has been changed to human clusters or settlements), Eeckhout (2004). In the meanwhile, a lot of studies try to find the economic explanations of Zipf's law which observed in cities in many countries. This also leads to further study about the determinants of urban growth. Specifically, Zipf's law states that city rank (ranked by population) is perfectly inversely proportionate to its size (refers to city population),$r(p) = A/p$, where r is the rank of city with population p and A is the population of the largest city (rank 1 city). In other words, within a country, the second largest city has the population about a half of the size of the largest, the third largest city has a third of the population of the largest, etc.

3.2.1 Pre-Zipf's Law

Zipf's law was first proposed by American linguist George Kingsley Zipf to address the fact that the frequency of any word is inversely proportional to its rank in the frequency table [23](Zipf 1935, 1949). This relationship has subsequently been found in many other areas like physical and social sciences such as the population ranks of cities in various countries, corporation sizes, income rankings and so on.

Actually, the first interesting empirical finding of this kind of phenomenon was proposed by Auerbach as early as 1913. He shows that in the U.S. and five European countries, city population consistently adhered to the relationship of

$$p_i r_i = A$$

, where p_i is the population of city i within a country; r_i is the rank (by population) of city i and A is constant. Then Lotka (1925) states that in the distribution of human

[23] For example, in the Brown Corpus, the word "the" is the most frequently occurring word, and by itself accounts for nearly 7% of all word occurrences (69,971 out of slightly over 1 million). True to Zipf's Law, the second ranked word "of" accounts for slightly over 3.5% of the number of whole words (36,411 occurrences), followed by "and" (28,852). Only 135 vocabulary items are needed to account for half the Brown Corpus.

agglomerations there appears to be a remarkable statistical regularity and proposes 'a parallel to Pareto's law' as the equation below to capture this regularity

$$logy = A - \alpha\ logx$$

which is exactly the equation used nowadays for testing for Zipf's law, when $\alpha = 1$, then Zipf's law holds. Then Lotka (1925) also tests this regularity using places' population data for several developed countries [24] and finds that the Pareto coefficient α is approximately 1.00 for U.S. (1920) and Germany (1925); 0.97 in England (1921) and 1.30 in France (1921). France is the exception as it is dominated by small towns. Moreover, Singer (1936) rewrites the relationship between rank and size as

$$rp_r^{\alpha} = A$$

or furthermore,

$$r = Ap_r^{-\alpha}$$

, where r is also the rank of a city by population within a country at a specific time. This is the normal equation of Pareto distribution with a power exponentα.

3.2.2 Zipf's Law and Theoretical Explanations

[24] Including U.S., Germany, France, England & Wales, Japan, Hungary and Canada.

In 1949, George Kingsley Zipf, a Harvard linguistics professor revisits the relationship between rank and size and states more specifically: if we rank city size by population in descending order from the largest (rank 1) to the smallest (rank N) to get the rank $r(p)$ for a city of size p, then:

$$r(p) = Ap^{-\alpha} \qquad (1)$$

where A and α are parameters, α is the so-called 'Pareto coefficient'. This is also called 'the rank-size rule' for cities. Zipf (1949) claims that not only does the size distribution of cities follow a Pareto law (expressed above), but also the distribution has a special shape that the parameter α equals to 1. Zipf (1949) shows that the rank of a city (ranked by population) is perfectly inversely proportional to its size (population).

$$r(p) = \frac{A}{p} \qquad (2)$$

Precisely, Zipf's law is a special case of the Pareto law or Pareto distribution which is a power law probability distribution that has been observed in real world from social, geophysical and many other science areas. The Pareto law is initially applied to the distribution of income, i.e. how many people having an income greater than x and given in terms of cumulative distribution function: if X is a random variable with a Pareto distribution, then the probability that X is greater than x is given by:

$$P(X > x) = \begin{cases} \left(\frac{x_m}{x}\right)^{\alpha} & \text{for } x \geq x_m \\ 1 & \text{for } x < x_m \end{cases}$$

where x_m is the minimum possible value of X, and α is a positive parameter, i.e.

$$P(X > x) \sim x^{\alpha}$$

Pareto's law shows that the probability of the number of events that are larger than x is an inverse power of x. Therefore, Zipf's law is a special case of the Pareto law/ Pareto distribution when the Pareto exponent equals to 1. Besides, in terms of continuity, Pareto distribution is a continuous probability distribution while Zipf's law, also sometimes called the zeta distribution, can be considered as a discrete counterpart of the Pareto distribution in city size distribution. Moreover, the Pareto law is a branch of Power law which is a mathematical relationship between two quantities where the frequency of some event varies as a power of some attribute of that event (e.g. size) as

$$\mathrm{f}\,(\mathrm{p}) = \mathrm{a}\, p^{\alpha}$$

where f (p) is the frequency of an event p.

Zipf's restatement evokes a large number of empirical studies attempting to test for the validity of Zipf's law. Empirical studies testing for Zipf's law usually test equation (3) below. Cities are ordered by population size, with the largest having rank 1, then we regress the logarithm of their rank on the logarithm of their size:

$$log\, r_i(p_i) = logA - \alpha\, log\, p_i + \varepsilon_i \qquad (3)$$

If Zipf's law holds, eq.(3) should yield a slope coefficient close to one, i.e., $\alpha = 1$.

Since Zipf's law has been proposed lots of literature has repeatedly confirmed it by empirical evidence both across countries and time periods (Allen, 1954; Rosen and Resnick, 1980; Guerin-Pace, 1995; Eaton and Eckstein, 1997; Gabaix 1999b; Brakman *et al*., 1999, 2001; Davis and Weinstein, 2002; Ioannides and Overman, 2003; Rose, 2005; Soo, 2007; Bosker, 2008; Jiang and Jia, 2010; Rozenfeld *et al*., 2009; Berry and Okulica-Kozaryn, 2012; Jiang and Liu, 2012; Giesen and Sudekum, 2010). However, from the end of the 1990s another concern about Zipf's law arose: this surprising empirical regularity needs a rigorous theoretical explanation as to why might city size distributions follow Zipf's law. The first attempt to build such an economic theory was Gabaix (1999) who was the first to explain Zipf's law using Gibrat's law- another well-known empirical regularity stating that city growth is independent of its initial size. Gabaix (1999) begins with estimating a Pareto coefficient equal to 1.005 using 135 largest metro areas for the U.S. in 1991, then proposes a model that for a fixed number of cities growing stochastically with the growth process being homogeneous with a common mean and variance, which is a reflection of Gibrat's law, then in steady state the growth process produces a city size distribution following Zipf's law with Pareto exponent equals to 1. In other words, models of city growth should deliver Zipf's law. However, city size processes must have the time to converge to Zipf's law. Monte-Carlo simulations show that seven decades are enough to reach the value of power exponent of 1 (ζ=1.05, and ζ=1.001 for twice that time). With respect to the number of cities, which is increasing over time at a rate ν, as long as the growth rate ν is not greater than the growth rate of existing cities γ: $\nu \leq \gamma$, then the steady state distribution still satisfies Zipf's law in upper tail. When $\nu > \gamma$, in the continuous-time case, the steady state distribution has an exponent ζ which is the positive root of $\zeta^2 - \left(1 - \frac{2\gamma}{\sigma^2}\right)\zeta - \frac{2\nu}{\sigma^2} = 0$, indicating the power exponent is greater than 1. For what type of urban growth can be identified to be with a common mean and variance? Gabaix supposes that shocks are *iid* and will affect population growth both positively and negatively.

However, some studies criticize Gabaix's work for lacking economic content; therefore studies trying to explain Zipf's law from different economic foundations have emerged. Duranton (2002, 2006 and 2007) finds that several economic mechanisms can derive the Zip's distribution pattern; in particular he finds that the churning of industries across cities can lead to Zipf's law. As a result, his focus was not on the exact shape of city size distribution, but instead to evaluate what the real drivers of urban growth and decline are. Eeckhout (2004) proposed an equilibrium of local externalities. The driving force of urban growth is assumed to rely on random local productivity and perfect mobility of workers, which will lead to log-normal distribution of cities, and he confirms that the population of all U.S. places (‘places’ used in Eeckhout’s study instead of ‘cities’) is distributed log-normally empirically using 25,359 legally bounded places of U.S. census data in year 2000 (including 1 person place). Rossi-Hansberg and Wright (2007) argue that cities emerge from an endogenous trade-off between agglomeration forces and congestion forces. The mobility of workers ensures the marginal product of labour is equal across cities, which is independent of city size, yielding constant returns to scale and balanced growth of cities. Therefore, a city size distribution is described by a power law with a power exponent equal to 1 (Zipf’s law). Benguigui and Blumenfeld-Lieberthatl (2007) use random multiplicative growth to explain Zipf’s law. Later, Cordoba (2008 a, b) claims that localization economy will generate Pareto city size distribution. The model generates Pareto distribution in a balanced growth path, in which all cities have the same expected growth rate[25]. Therefore the standard model generates a Pareto city size distribution when (a) preferences for goods follows random walks and the elasticity of substitution between goods is 1; (b) total factor productivity of different goods follows random walks and increasing returns are equal across goods. Semboloni (2008) proves that Zipf’s law comes from the asymmetrical exchanges among cities. To summarize, basically all the theoretical models that try to explain

[25] Which can be achieved when the economy under one of the three conditions: (a) the elasticity of substitution between good is 1; (b) externalities are equal across goods; (c) critical conditions on preferences and technologies are satisfied.

Zipf's law rely on cities' local externalities which must be randomly distributed and independent of city size.

3.2.3 Debate on Testing for Zipf's Law

Since Zipf's law was launched numerous studies try to test for the validity of Zipf's law either in cross countries or within a country. Most of the empirical work is concentrated in developed countries, especially U.S., and finds empirical evidence for Zipf's law both cross-countries and over time. Eeckhout (2004) questioned whether Pareto distribution or log-normal distribution can approximate the city size distribution better. Eeckhout (2004) asserts that when we consider all cities in the U.S. it displays a log-normal distribution. Then in the last decade, there is a huge debate about many other aspects related to testing for the validity of Zipf's law, like the testing method or the definition of city or whether Zipf's law is a theoretical economic phenomenon.

Basically, there are three branches in the debate, based on the discussion of Berry and Okulica-Kozaryn (2012): (1) the first group is in a competing debate about whether a Pareto distribution or a log-normal distribution can approximate the city size distribution better and even more statistical distributions are proposed, or whether the urban growth paths are consistent with Gibrat's law over time. Furthermore the debate extends to what are the driven forces or determinants of urban growth (Parr and Suzuki, 1973; Krugman, 1996; Malacarne *et al*., 2001; Ioannides and Overman, 2003; Black and Handerson, 2003; Eeckhout, 2004, 2009; Garmestani *et al*., 2007; Soo, 2007; Levy, 2009; Giesen *et al.*, 2010; Rozenfeld *et al*., 2009; Berry and Okulica-Kozaryn, 2012; Michaels *et al*., 2008). (2) The second group of scholars is concerned with the methodology and the definition of a city, and even question whether Zipf's law is an economic regularity or it is merely a statistical phenomenon. Some researchers argue that the OLS method involved in testing for Zipf's law might be an inappropriate estimator for the Pareto exponent, given that the dependent variable-rank- is an integer and the intercept is not a nuisance parameter in the regression eq. (3) (Urzua, 2000 and 2011; Gabaix and Ibragimov, 2011; Rozenfeld, 2009; Jiang and Jia, 2010; Jiang and Liu, 2012; Michaels *et al.*, 2008). (3) In the meanwhile, some studies even doubt whether we really need an economic theory to support Zipf's law, because the skewed distribution functions of city size are uniquely stochastic steady states or Zipf's law is merely a statistical phenomenon (Axtell and Florida, 2001; Semboloni, 2008; Gan *et al.*, 2006; Batty, 2006 and Corominas-Murtra and Sole, 2010).

3.2.3.1 Pareto Distribution or Log-normal Distribution

Firstly, the surprising regularity-Zipf's law- has been repeatedly confirmed both cross countries and during different time periods. For cross-country studies, Rosen and Resnick (1980), Parr (1985), Soo (2005) and Rose (2005) provide the most complete international comparative studies and confirm the existence of Zipf's law. The classic international test of Zipf's law is Rosen and Resnick (1980)'s study, focusing on 44 countries in 1970 and they find that Zipf's law holds in 2/3 countries with OLS estimated Pareto exponent of α ranges from 0.81 to 1.96, and an average of 1.13 with a standard deviation of 0.19. While 1/3 of the countries' cities have more even distribution than Zipf's law predicts (the Pareto exponent is much higher than 1, the size difference between large cities and small cities is lesser). They also show that results are sensitive to data definitions that Pareto exponent is close to one when using agglomeration data rather than administratively defined 'cities'. Parr (1985) applies the Pareto law of income distribution to the city size distribution based on 12 countries and argues that within a nation the evolution of a nation's Pareto coefficient over time tends to display a U-shaped pattern. And the position of a nation in its sequence will depend on the overall level of development or perhaps its age. According to Singer (1936), the Pareto exponent can be considered as an index of urbanization or a measure of city size inequality which can give a measure of the proportion of large cities or small cities, i.e. the lower the value of the Pareto coefficient, the higher the proportion of large cities. Then Soo (2005) studies more comprehensively 73 countries' census data from 1972 to 2001 and confirms that Zipf's law exists but is rejected far more often than we would expect. For 53 out of 73 countries Zipf's law is rejected using OLS method. The largest Pareto exponent is 1.719 found in Kuwait followed by Belgium with a Pareto exponent 1.5895, while the lowest Pareto exponent is found in Guatemala at 0.7287, which is consistent with the reality that Kuwait and Belgium both have a large number of small cities with no primate city (higher Pareto exponent expected, the slope of plotting log rank on log size is relatively large). While 30 out of 73 countries reject Zipf's law using Hill estimator. Rose (2005) considers Zipf's law and Gibrat's law extending to country-level instead of city-level, i.e. country size distribution can also follow a Zipf's law or Gibrat's law. He studies 50 largest countries for Zipf's law and 163 sovereign countries for Gibrat's law from 1900 to 2005[26], finding that Zipf's law holds using conventional OLS method- regressing log rank on log size. He asserts that none of the Pareto exponents are different from 1 at traditional confidence levels for these years (the biggest exception is 1900 whose Pareto exponent is slightly over one standard deviation from 1). For Gibrat's law, using OLS method, regressing log population growth between 1990 and 2000 on log 1990's population, he finds Gibrat's law holds in country size i.e. that the growth of country size is independent of its initial size.

The majority of empirical studies focus on whether Zipf's law holds in cities within a country. Krugman (1996) confirms Zipf's law using U.S. metropolitan areas in 1991 with a Pareto exponent exactly equal to 1.005. Eaton and Eckstein (1997) find evidence for France and Japan's top 40 cities from 1876 to 1990. They find for both

---

[26] They use data for 1900, 1950, 1960, 1970, 1980, 1990,2000,2004,2005.

countries city size distribution can be described quite well by 'rank-size rule' (Zipf's law) using OLS estimation for Pareto exponent: for France the Pareto exponent equals 1 except for early period 1875 as 0.87 (more unequal); while for Japan, the Pareto exponent is slightly less than 1 and lower in 1925, 1947 and 1950, indicating greater inequality for city size in Japan comparing to France. They also support Gibrat's law by finding no difference in the mean growth rate between large cities and small cities. Gabaix (1999a) explains Zipf's law using Gibrat's law for the first time and also tests Zipf's law using U.S. data for 135 large metro areas in 1991. He firstly proposes a model showing that after homogeneous growth in cities, in steady state, the distribution of cities will follow Zipf's law with a power exponent equal to 1, and then finds a Pareto exponent equals to 1.005 for U.S. 135 large metro areas in 1991. The growing number of cities over time is also considered in his paper, he shows that as long as the appearance rate of new cities is no greater than the growth rate of existing cities, the Pareto exponent will not be different from 1 (Zipf's law holds). Besides, he also finds Gibrat' law holds, i.e. cities follow a homogeneous growth process with a common mean and a common variance. Dobkins and Ioannides (2000) use a longer time span of U.S. metropolitans population (census data) from 1900 to 1990 and conventional OLS and Maximum likelihood methods to estimate the Pareto exponent. They confirm the validity of Zipf's law by finding the Pareto exponent clearly close to1: 1.044 in 1900 and 0.949 in 1990 from OLS method and 0.953 and 0.553 in 1900 and 1990 respectively from Maximum likelihood method. Davis and Weinstein (2002) study for a long period for Japan from Stone Age to modern era and find that the distribution of regional population is very close to Zipf's law in Jomon and Yayoi (-300 to 300) periods when primitive agriculture and ethnically Japanese people firstly appear, with some metallurgical skills, some coins but no writing or cloth. This study might confirm that the distribution of population agglomeration tends to display Zipf's law under stochastic population growth, i.e. Gibrat's law can generate Zipf's law. Ioannides and Overman (2003) use the same data as Dobkins and Ioannides (2000), U.S. census metropolitans' population from 1900 to 1990 to test for the validity of Zipf's law. They calculate 'local Pareto exponents' from the mean and variance of city growth rates and find that Zipf's law is broadly satisfied for most of the samples, but may vary according to city size. Using non-parametrical estimation of a stochastic kernel, a three dimensional representation of the distribution of growth rate conditional on city size, they find Gibrat's law does hold in urban growth, and any deviations from Zipf's law can be explained by deviations from Gibrat's law.

However, other researchers find Zipf's law does not hold well. Black and Henderson (2003) use the same period U.S. data as Dobkins and Ioannides (2000), Ioannides and Overman (2003) from 1900 to 1990 but construct more consistent metropolitan data during decades. In contrast to previous findings, they show that Zipf's law only holds for upper third cities (largest one third) and for the full sample the Pareto coefficient is around 0.85 for any decade, using conventional OLS methods to estimate the Pareto exponent. The fact that the Pareto coefficient in any decade is much greater for the top one-third of cities than for the whole sample indicates the relationship between rank and size is not log-linear, and a quadratic term has been found in the regression throughout the whole period. Besides, Gibrat's law is rejected for any sample size, by regressing the log subsequent population growth on log initial size. Black and Henderson's finding highlights the point that results may be extremely sensitive to the geographical unit chosen (city definition) and sample size. Furthermore, Eeckhout (2004) proves that if we consider all U.S. places city size distribution follows a log-normal distribution not Pareto, using census 2000 population data for all places in U.S. (25,359 legally bounded places, even including 1 inhabitant place). He emphasizes the high sensitivity of the Pareto coefficient to the sample size, and argues that the conventional OLS procedure of testing for the Pareto coefficient is not equivalent to the goodness-of-fit test for whether Pareto distribution can approximate the city size distribution. Therefore he uses a Kolmogorov-Smirnov (KS) test of goodness-of-fit of the empirical distribution against the theoretical distribution and finds that log-normal fits the whole data quite well. His explanation for why Zipf's law repeatedly holds while the entire underlying distribution of cities is log-normal is that the density function of the Pareto distribution is not dramatically different from the density function of log-normal at the very upper tail. Previous studies confirming Zipf's law are using an upper truncated sample, for instance, 135 metropolitan areas in U.S. cities. In other words, the Pareto distribution can be a statistical phenomenon in the upper tail of the log-normal distribution. Then some studies find Zipf's law or Pareto distribution does not fit the city size data well. Garmestani *et al.* (2007) note deviations from Zipf's law using U.S. south-eastern region cities from 1860 to 1990 (by decade). Simulations are established by calculating a kernel density estimate of the log-transformed data (Hall and York, 2001) and compare with the actual data with a null distribution, they find that city size distribution of U.S. south-eastern region follow a log-normal distribution. However, Gibrat's law is rejected using hypothesis testing and graph analysis. City growth is correlated to size, small cities have higher growth rate while large cities have lower growth rate in south-eastern U.S., indicating urban hierarchy is discontinuous. Soo (2007) studies Malaysian cities with five censuses on 1957, 1970, 1980, 1991, 2000. Zipf's law is rejected for full sample for all periods (except 1957), but approximately fits the data at upper tail. However, Gibrat's law is rejected by finding evidence that smaller cities grow relatively faster, as well as state capitals and cities in the states of Sabah and Selangor.

Levy (2009) raises doubts about the fit of very upper tail of the city size distribution in Eeckhout's (2004) work. Specifically, he is unsure about whether the upper tail follows log-normal, because he finds evidence that the very upper tail still follows a Pareto distribution. He uses the same data as Eeckhout (2004), U.S. 25,359 places population, and finds that Zipf's law holds in the upper tail but fails in the bulk of the distribution; while log-normal distribution is rejected for large cities in the upper tail by a graphical analysis. Then other works reconfirm Zipf's law and accordingly Pareto distribution. Bosker et al. (2008) find Zipf's law holds for 62 West-German cities before WWII. The WWII shocks the city size distribution from one adhering closely to Zipf's law to one characterized by a more equal city size distribution

(Pareto coefficient increasing). Berry and Okulica-Kozaryn (2012) reconfirm Zipf's law holds quite well using U.S. Economic Areas data from 1990 to 2010. The OLS estimated Pareto coefficients are 1.009, 0.994, and 0.986 for 1990, 2000 and 2010 respectively. Gibrat's law also holds well by regressing the log of the current population on log previous population, hoping to find the estimated coefficient close to 1 if Gibrat's law is expected to hold. They find the estimated coefficient is 1.004 for 1990 to 2000 and 1.015 for period 2000-2010.

3.2.3.2 Definition of Cities, Testing Methodology and Doubts about Zipf's Law

The debate has attracted more attention in recent years, especially in the city definition aspect. The city definition problem is often considered as one of the reasons for deviations of Zipf's law, as cities are defined by administrative boundaries not population clusters or economic agglomerations. Due to the geographical technology development, recently, researchers testing for these two laws use more delicately defined cities, such as Jiang and Jia (2010) using U.S. 'natural cities' which are constructed by clustering street nodes observed by satellite to test for Zipf's law. They find that Zipf's law holds remarkably well for all natural cities (over 2.4 million natural cities in total), as expected. Then more powerful evidence is found by Rozenfeld *et al*. (2009) who build cities 'from the bottom up' by clustering populated areas obtained from high-resolution data and form U.S. and UK's population agglomerations data. Using these city agglomeration data, they reconfirm Zipf's law for the whole sample, including cities as small as 12,000 inhabitants in U.S. and 5,000 inhabitants in UK. Jiang and Liu (2012) define city boundaries by grouping smaller blocks for France, Germany and UK. They find city size distribution indeed exhibits a power law distribution $P(x) \sim x^{-a}$. Moreover, Michaels *et al.* (2008) using U.S. sub-county data Minor Civil Divisions that cover the rural area (but just for the Midwest and Northeast area, others using county data), for 1880, 1940 and 2000, find that Gibrat's law is strongly rejected when both rural and urban areas are considered. For medium density counties there is a positive relationship between population growth and initial size. As mentioned before, the explanation of Zipf's law is Gibrat's law; therefore the deviation of Gibrat's law will lead to the deviation of Zipf's law. Therefore, according to Michaels *et al.*'s study Zipf's law does not hold when including rural areas.

With respect to the testing method for Zipf's law, it has been customary to test the law by simply plotting the logarithm of rank against the logarithm of size hoping to find a straight line with a slope of minus one, i.e. the 'Zipf's plot' as in section 3 and 4. Then more formally, through OLS regression of

$$log\, r_i(p_i) = logA - \alpha\, log\, p_i + \varepsilon_i$$

as mentioned before, testing whether the estimated coefficient of α is close to -1. Urzua (2000, 2011) criticize this traditional testing procedure for Zipf's law as being inefficient, since rank r_i is an integer, the distribution of ε_i is far from being normal. Urzua proposes a more simple and efficient Lagrange multiplier (LM) test as

$$\mathrm{LMZ} = 4\mathrm{n}({z_1}^2 + 6z_1z_2 + 12{z_2}^2)$$

where

$$z_1 \equiv 1 - \frac{1}{n}\sum_{i=1}^{n} ln\frac{x_i}{x_{(n)}}$$

and

$$z_2 \equiv \frac{1}{2} - \frac{1}{n}\sum_{i=1}^{n}\frac{x_{(n)}}{x_i}$$

where x_i is the population of city i, and n is the total number of cities. For instance, Urzua test the *LMZ* value for 135 U.S. metropolitan areas in 1991, as $x_{(135)} = 252{,}000$ (the population of Charleston, WV) the calculated resulting value for LMZ is 3.16. Thus according to the significance points for LMZ (Urzua's Monte Carlo simulations), the null hypothesis that $\alpha = 1$ cannot be rejected at 10% significance level, which is also consistent with the study of Krugman (1996) and Gabaix (1998) using the same data. Besides, the other prevalent method improvement is made by Gabaix and Ibragimoy (2011) who propose that adding a shift of 0.5 for the rank in the conventional OLS regression equation testing for Zipf's law is optimal and can avoid the bias of OLS estimation in small samples. They modify the regression equation as

$$\ln(Rank_i - 1/2) = lnA - \alpha \ln pop_i + \varepsilon_i$$

for estimations for small samples. Studies using these different methods may obtain a different results and different Zipf's exponents.

During the debate of whether Zipf's law holds, there are some other researchers who doubt that the relationship indicated by Zipf's law holds. Semboloni (2008) proposes a stochastical dynamic model showing that a hierarchy arising from bottom-up mechanism and resulting the power law distribution of cities. Gan et al. (2006) argue that Zipf's law is spurious in explaning city size distribution as it is merely a statistical phenomenon, not suggesting an economic regularity. They use Monte Carlo simulations to examine the conventional rank-size regression equation by running on random numbers (and their ranks) which generated from various probability distributions. Corominas-Murtra and Sole (2010) consider that city size or firm size distributions are just examples of the universal law of Zipf's. They explain Zipf's law by assuming that the complexity of the distribution system provided by the sequence of observations is the one expected for a system evolving to a stable state between order and disorder, the result is obtained from several assumptions, models not involved. Corominas-Murtra and Sole (2010) claim that the general nature of their derivation and the model-free basis would explain the universality of Zipf's law in real systems.

3.2.3.3 Deviations from Zipf's Law

If we accept that Zipf's law cannot fully approximate the city size distribution, researchers further explore the explanation and measurement of deviations from Zipf's law. Theoretically, as mentioned before, most of the literature believes that Zipf's law is delivered by Gibrat's law, and a deviation from Gibrat's law will lead to the deviation of Zipf's law. The explanations for smaller cities' often having a relatively smaller Pareto exponent is that the variance of their growth rate is larger empirically, since deviations from Zipf's law are found in different countries (relative literature is mentioned above in the debate that against Zipf's law), researchers propose some other statistical distributions that can approximate city size distribution instead of Pareto distribution, like log-normal distribution (Parr and Suzuki, 1973; Eeckhout, 2004), q-exponential distribution (Malacarne *et al*., 2001; Soo, 2007) or double Pareto log-normal distribution (Giesen et al., 2010). Soo (2005) not only finds 53 counties out of 73 countries reject Zipf's law, as mentioned above, but also attempts to explain the variation in Pareto exponents in different countries using independent variable as income per capita, transport costs, population, public spending and political variables and dependent variable as Pareto's exponent per country. Soo's (2005) results show that the Pareto exponent is positively related to a country's per capita GNP, total population and railroad density; negatively related to land area. Gonzalez-Val (2011) measures deviations from Zipf's law of 23,519 places in U.S. in 2000 census data (the same data as Eeckhout, 2004), by regressing the population deviations from Zipf's law on city characteristics. Gonzalez-Val finds 18,874 places (80.25%) show negative deviations indicating that population is higher than Zipf's law prediction; 4645 places have positive deviations indicating the size is lower than Zipf's prediction. The most important determinants of a city presenting a deviation are per capita income, human capital levels and the proportion of population employed in some sectors.

Furthermore, to study whether Zipf's law or Gibrat's law holds and the city size distribution, some studies examine the determinants of urban growth and urban growth theory since urban growth patterns will affect Gibrat's law and further Zipf's law and shape the distribution of city sizes. For instance, Glaeser et al. (1995) examine the growth pattern of the 200 most populous cities in the U.S. from 1960 to 1990 and find that the growth relies on various initial characteristics of cities in 1960. They find that city population growth and income growth are positively correlated with a city's initial education level and negatively related to initial unemployment and initial percentage of employment in manufacturing industries. Black and Henderson (1999) explore how urbanization affects economic growth and how growth affects patterns of urbanization. They raise doubts about the validity of Zipf's law as an empirical regularity and construct a theoretical model to explain the parallel growth of cities with a Markov chain. Also, empirically they find a positive relationship between growth in city sizes and growth in local human capital levels using 318 U.S. metropolitans' population data from 1940 to 1990 by decade. Glaeser and Shapiro (2003) study the determinants of city growth using U.S. cities (with inhabitants over 25,000 and MAs) from 1990 to 2000. They employ a wide range of explanatory variables of initial city characteristics and find the three most influential variables on city growth are human capital, climate and transport systems.

To conclude, in general many empirical studies show that Zipf's law for cities is often found in the upper tail of the sample, especially in the U.S. and other developed countries, i.e. Zipf's law holds at least for large cities. Then many researchers attempt to explain this surprising regularity and the most prevalent explanation is that Gibrat's law could deliver Zipf's law. Basically all the theoretical models that trying to explain Zipf's law rely on city's local externalities which must be randomly distributed and independent of city size. After Gibrat's law is considered, there is a debate about whether Zipf's law representing Pareto distribution or Gibrat's law representing log-normal distribution could approximate the real city size distribution. Then Eeckhout (2004) argues that both laws are consistent because a lognormal upper tail can typically not be distinguished from a Pareto (Zipf) upper tail, the plot of density functions of lognormal and Pareto distribution shows that the Pareto distribution is very different from lognormal, however, at the very upper tail of the distribution there is no dramatic difference between the density function of the lognormal and Pareto.

This is the case in most studies as they use a truncated sample of large cities. As a result, both the Pareto and truncated lognormal distribution match the data relatively closely. The problem is that the estimated Pareto coefficient is extremely sensitive to the choice of the truncation point, for lower truncation points, the Pareto fits the data less and less well, and also the definition of cities and testing method for Zipf's law can affect results.

3.2.4 Empirical Evidence in China

The empirical literature applying Zipf's law to China and other developing countries' cities is relatively scarce. The majority of the research supports Zipf's law for Chinese cities at the upper tail of large cities (Song and Zhang, 2002; Gan, Li and Song, 2006; Soo, 2014; Schaffar *et al*., 2010; Ye and Xie, 2012). However, the data used in these studies is insufficient, merely 2 or 3 years repeated cross-sectional regression. Therefore, some recent studies reject Zipf's law for Chinese cities using different years and methods (Anderson and Ge, 2005; Xu & Zhu, 2009; Peng, 2010; Ye and Xie, 2012). Nonetheless, all the relative studies on Chinese cities are not comprehensive as they are testing the city size distribution only on certain years and conventional or modified OLS methods (which has its weakness for integer dependent variable- rank-as mentioned before).

Some studies support for Zipf's law. The earliest study on Chinese city size distribution, Song and Zhang (2002), consider city-level (county-level cities in China) data in 1991 (479 cities) and 1998 (665 cities). Using the non-agriculture population of cities and conventional OLS method of testing for Pareto exponent, they estimate that Pareto coefficient to be 0.92 in 1991 and 1.04 in 1998 (within the range that support Zipf's law). The variation of the value of the Pareto exponent indicates that Chinese city size became more even in the 1990s, from 0.92 (a little bit less even than Zipf's law predicts) to 1.04 very close to Zipf's law perfectly evenly distribution. In addition, they also test whether the Pareto exponent is sensitive to sample thresholds and size; results show that it is sensitive and that the Pareto exponent increases with a higher cut-off threshold, in a large city sample, the Pareto exponent increases to 1.39 when 210,000 inhabitants is the threshold and sample size is 271 cities). This contradicts the empirical findings for the Zipf's exponent in U.S. cities claiming that Zipf's law holds at least in upper truncated samples. Moreover, they added a quadratic term and find non-linear relationship between rank and size, and the R^2 rising shows that the quadratic model fits the actual distributions better than a Pareto law. Gan, Li and Song (2006) also test Zipf's law using 1985 and 1999 city level data, and suggest that Zipf's law fits well for city-size distributions in China with Zipf's law coefficients of 0.86 and 1.08 for 1985 and 1999, respectively. However, they find that the data does not imply a Pareto distribution through a Kolmogorov-Smirnov nonparametric test. The K-S test checks the equality of distributions by comparing the empirical distribution of the data with a given theoretical distribution, and finds that both the 1985 and 1999 data reject both Pareto distribution and log-normal distribution with p values equal to 0. This suggests that the city-size distribution in China does not follow a Pareto distribution even though the Pareto exponent is close to1. Soo (2014) supports Zipf's law using provincial level census data for 1953, 1964, 1982, 1990, and 2000. He tests Zipf's coefficient in 1953 and 2000 using a modified OLS rank-size regression (Gabaix and Ibragimov, 2011), and find that the Zipf's coefficient decreases from 3 when the sample size is 10 provinces, to about 0.8 when all provinces are included. Zipf's law is rejected only for the sample of 20 provinces in 2000. By comparison, the average Zipf coefficient for cities around the world is about 1.1 (Nitsch, 2005; Soo, 2005). Soo (2014) also investigates Gibrat's law by System GMM method using the following estimating equation:

$$\ln(P_{i,t}) = \beta + \theta_i + \gamma_t + \delta \ln(P_{i,t-1}) + \varepsilon_{i,t}$$

to see whether $\delta = 1$. He rejects the null hypothesis that Gibrat's law holds, controlling for the endogeneity of population. He suggests that in Chinese provinces, population growth is influenced by initial population levels. Schaffar *et al*. (2010) firstly provides a Kolmogorov-Smirnov (KS) test for Chinese cities (225-577 cities) for 1984, 1994 and 2004, showing that the Chinese city-size distributions for cities over 100,000 inhabitants follow a Pareto law. Secondly, the same K-S test for lognormal distribution found that the null hypothesis is always rejected, i.e. Gibrat's law is rejected in China. Thirdly, to explore the urban growth patterns they run the second-generation panel unit root test and found that a unit root cannot be rejected which means that the underlying trend is stochastic and there is no steady state size: city sizes do not converge over time in China (city size evolve in a non-stationary way). In the 1980s small-sized cities grew faster, and in the 1990s, higher growth was seen for medium-sized cities. In addition, they test for the co-integration to see if there is parallel growth of cities within the same province. Contrary to our expectations, the geographical location of a province does not seem to play a significant role in the appearance of parallel growth; in this sense urban growth patterns do not reject Gibrat's law for cities in China.

In contrast, a few empirical studies reject Zipf's law for Chinese cities. Anderson and Ge (2005) find that various power laws are strongly rejected for Chinese cities while Gibrat's law does describe the Chinese situation well when they are using city level data (77-658 cities) from 1949 to 1999 (1949, 1961, 1970, 1980, 1985, 1990, 1994, 1999). Specifically, firstly using standard OLS methods, they perform a panel data regression for Zipf coefficient and then test for the year dummies for these 8 years, finding that for the period 1949-1980 (before the 'Economic Reform' period) the estimated θ is not significant, while from 1980 to 1999 the Pareto exponents are significantly higher than 1, implying that the city size distribution in China is more even than Zipf's law would predict. Secondly, through efficient OLS together with maximum likelihood of the Pareto coefficients together with the corresponding goodness of fit tests (Pearson, 1900) they found that the Pareto coefficients are lower in the standard OLS estimations, and still the Zipf's law is rejected, except for 1949. Using the same method they find that the log-normal distribution cannot be rejected, which is supportive of Gibrat's law (for cities with a size above 100,000 inhabitants). Anderson and Ge's (2005) contrary results may be because of the different data and method. City size data are using urban agglomeration and the testing method is to estimate the Pareto exponent using panel data and the maximum likelihood method and corresponding Pearson tests. They adopt the city criteria based on 1963[27] which means in fact they are using the urban agglomeration data rather than city-proper data as in this paper. As mentioned before, for Zipf's law, it does matter whether one deals with urban agglomerations (i.e. metropolitan areas) or with city-proper data (Gabaix and Ioannides, 2004). The Pareto exponent should be larger for city proper than the urban agglomeration data, because urban agglomerations are not bound by legal definitions, unlike city-proper and therefore they are likely to have a longer upper tail. This point was made first by Rosen and Resnick (1980) and has been revisited by Brakman et al. (1999, 2001). Furthermore, Xu and Zhu (2009) reject both Zipf's law and Gibrat's law by studying the growth trend of cities of different sizes of China from 1990 to 2000. Using non-agriculture city population, their OLS estimated Pareto exponents suggest that city size has become more evenly distributed, which implies a decreasing urban concentration. Besides, they also confirm the persistent convergence tendency in Chinese city size growth processes in 1990s, regardless of different urban population definitions (non-agriculture population or total city population), sample

[27] According to Chinese city criteria in 1963, city is defined as an urban agglomeration with a total urban population larger than 100,000 inhabitants.

divisions (all cities, 441 cities in total or prefecture-level cities, 166 cities in total) or estimation models (absolute convergence or conditional convergence which includes control variables for city characteristics). That means in 1990s small or medium-sized cities are growing faster than large cities. Peng (2010) finds deviations from Zipf's law for Chinese cities from 1999 to 2004 (every year) but the Pareto exponent is not so far from 1 (mean 0.84), using a new method of 'rolling sample regressions'. The sample is changing with truncation point, if Zipf's law holds with a Pareto exponent equals to 1, rolling sample regressions should yield a constant coefficient regardless of sample sizes. He finds that the Pareto exponent is almost monotonically decreasing with the rolling from rank top 100 to rank top 500 and to full sample, from around 1.8 to around 1.2. Ye and Xie (2012) still do not support Zipf's law using more detailed Chinese cities. They re-examine Zipf's law for Chinese cities in a more detailed manner, using Zipf's plots for different regions' cities and top 100 Chinese cities respectively from 1960 to 2000 (by decades). They divide China into six regions, see Figure 3.1 below, based on geographic positions (east, central-south, north, northeast, northwest, and southwest, Xie and Dutt, 1990) and find that none of the regions' cities follow the prediction of Zipf's law. Specifically, East and Northeast regions' city size distribution is much more even than Zipf's law prediction; North and Northwest seem to be closer to Zip's law while Southwest and Central-south is still in transition to be either closer to Zipf's law or contrary to Zipf's law because cities are constantly changing their ranks. For the top 100 cities from 1960 to 2000, Pareto exponents are increasing from 1.39 to 1.87 [28] indicating much more even city size distribution than Zipf's predicts. While for the whole sample, the Pareto exponent increases from 1.16 to 1.34 during 1960 to 2000, which still do not follow Zipf's law.

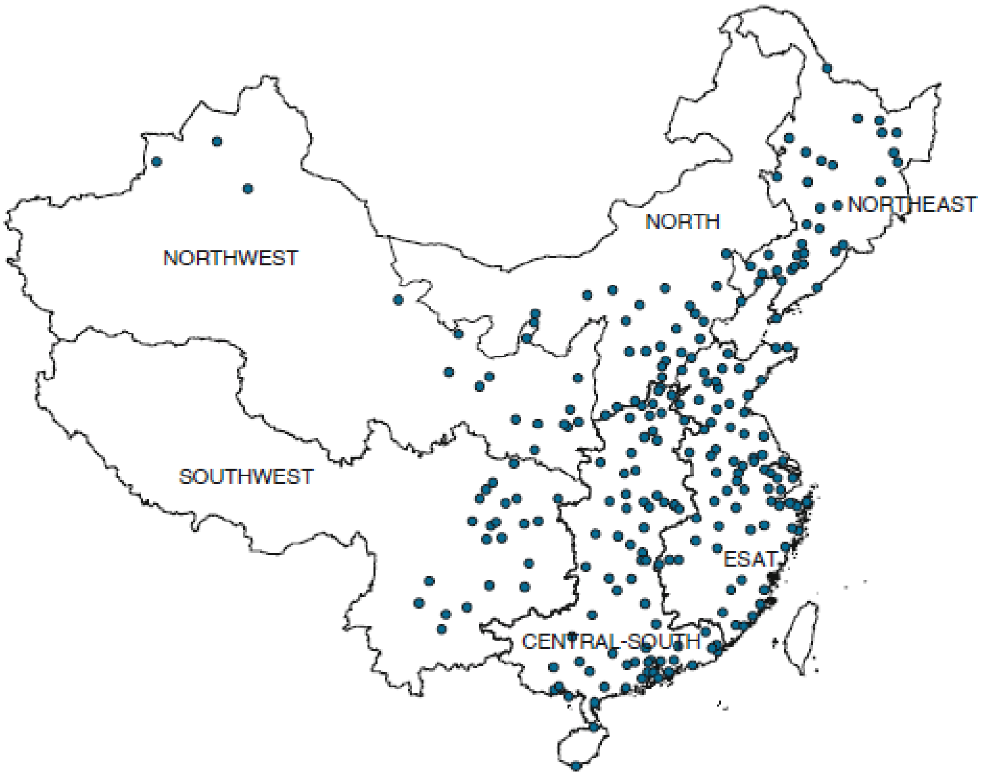


Figure 3.1 Six Chinese regions, divided in Ye and Xie's (2012) paper.

In conclusion, Zipf's law is also debatable in Chinese cities. On one hand most empirical investigations for Chinese city size distribution generally support Zipf's law (but only for a few specific years). On the other hand, there are a few exceptions, like Anderson and Ge (2005) for years 1949 to 1999 and Xu and Zhu (2009) for 1990 to 2000, using different urban population data (urban agglomeration and non-agriculture

[28] Ye & Xie's (2012) paper uses log population as dependent variable and log rank as independent variable, unlike most of the literature's customer, therefore the Pareto exponent is transformed by ourselves according to Ye & Xie's results: Pareto exponent = 1/0.7183=1.39 for 1960 and 1/ 0.534=1.87 for 2000.

population respectively, one could argue is not suitable for Chinese cities, explained in more detail in section 3.1) and different estimation methods.

3.3 Methodology

To begin with, testing for Zipf's law, in general, two types of methods have been used in the literature: graphs and regressions. Begin by ranking cities by the size of their population, descending, and the largest city is denoted rank (Chongqing is currently rank1 city in China, with population 15,427,700, Shanghai rank2 and so forth). One then compare the natural logarithm of city rank to the natural logarithm of city population, using either a graphical method or regression techniques. Figure 3.2- Zipf plots presents a set of graph that scatters the logarithm of city rank (by population) against the logarithm of city size (population) for each year, from 1910s to 2008, in order to illustrate Zipf's law.

For the regression part, the most intuitive and commonly used method in the literature is OLS estimation; accordingly we estimate the following three regressions:

$$\mathrm{logrank}_i = (\mathrm{logA})_1 + \alpha_1 \mathrm{logpop}_i \tag{4}$$

$$\mathrm{logrank}_i = (\mathrm{logA})_2 + \alpha_2 \mathrm{logpop}_i + \beta_2 (log\mathrm{pop}_i)^2 \tag{5}$$

$$\text{logrank}_i = (\text{logA})_3 + \alpha_3 \text{logpop}_i + \beta_3 (log\text{pop}_i)^2 + \gamma_3 (log\text{pop}_i)^3 \qquad (6)$$

eq. (4) seeks to test whether $\alpha_1 = -1$ and A=size of largest city, while eq. (5) and (6) aim to uncover any nonlinearities that could indicate deviations from Pareto distribution. Literature shows only quadratic term and cubic term of logpop_i, no evidence show that furture powers do not matter. Thus, one can attempt to add in more power of logpop_i to the regression equation. All of these regressions are run for each year, i.e. repeated cross sectional regression (because the rank variable is only valid for each individual year), using OLS with heteroskedasticity-robust standard errors. For each sample (full sample, administrative sample, regions sample, and historical sample) we estimate these three regressions respectively.

One potentially serious problem with the Zipf regression is that it is biased in small samples. Gabaix and Ioannides (2004) show using Monte Carlo simulations that the coefficient of OLS regression of Eq. (4) is biased downward for sample sizes in the range that is usually considered for city distribution. Thus the Gabaix and Ibragimov (2011) correction is used which considerably reduces the bias for the OLS method when applied to small finite samples:

$$\ln(Rank_i - \delta) = lnA - \alpha_4 \ln pop_i$$

With $0 \leq \delta < 1$. When $\delta = 0$, Zipf's law holds perfectly. According to Gabaix and Ibragimov (2011), the best estimation of α_4, for small samples, is provided when $\delta = 1/2$, thus:

$$\ln(Rank_i - 1/2) = lnA - \alpha_4 \ln pop_i \qquad (7)$$

With the most closet to true standard error of the estimated coefficient given by (Kratz and Resnick, 1996; Gabaix and Ibragimov, 2011; Dimou and Schaffar, 2009):$\delta \sqrt{2/n}$.

In addition, we employ the new Lagrange Multiplier (LM) parametric method to test for Zipf's law in Chinese cities according to Urzua (2011), to avoid the common pitfall in testing for Zipf's law. Urzua (2011) points that not only the fact that OLS estimators are not efficient as mentioned above, but it is also plainly wrong because the intercept is not a nuisance parameter in the regression since the dependent variable rank ($\boldsymbol{Rank_i}$) is an integer.

The procedure of producing LMZ statistics is as follows: chose $\boldsymbol{n}$ cities within a country and rank them by size (urban population) to get the ordered sequence

$$\boldsymbol{x}_{(1)} \geq \boldsymbol{x}_{(2)} \geq \cdots \geq \boldsymbol{x}_{(r)} \geq \cdots \geq \boldsymbol{x}_{(n)}$$

Zipf's law, known as rank-size relation, asserts that a graph of the rank against the size would then render a rectangular hyperbola. That is,

$$\mathbf{r}\,\boldsymbol{x}_{(r)} = \boldsymbol{c}$$

for a constant $\boldsymbol{c}$ and all $\mathbf{r}$.

or

$$\mathbf{r} = \mathbf{c}\,\boldsymbol{x}_{(r)}^{-1} \qquad (8)$$

Most of the empirical studies that verify Zipf's law use the erroneous procedure, they estimate, through OLS like Eq. (4) above for $\mathbf{r} = \mathbf{1}, \ldots, \mathbf{n}$, and then test the null hypothesis that$\boldsymbol{\alpha}_{\mathbf{1}} = -\mathbf{1}$. Aside from the fact that the OLS estimators are not efficient in that case (since $\boldsymbol{r}$ is an integer, the distribution of $\boldsymbol{\varepsilon}$ is far from being normal), given that $\boldsymbol{r}$ is an integer, what makes Eq. (4) incorrect is the fact that the intercept is not a nuisance parameter in the regression.

As it has been forcefully noted by several authors over the years (e.g. Quandt, 1964; Rapoport, 1978; Kamecke, 1990), before testing for Zipf's law one has to make explicit the underlying probabilistic process. As shown by, e.g., Rapoport (1978) or Urzua (2000), the probability law behind Eq. (8) $\mathbf{r}\,\boldsymbol{x}_{(r)} = \boldsymbol{c}$ corresponds to the Paretian density function $\mathbf{f(x)} \propto (\frac{x}{\mu})^{-2}$, where $\mathbf{x} \geq \boldsymbol{\mu} > \mathbf{0}$. Among a number of parametric and nonparametric tests for Pareto distribution, Urzua (2000) proposes in particular the following simple test statistic, the Lagrange multiplier (LM) test for Zipf's law,

$$\mathbf{LMZ} = \mathbf{4n}(z_1{}^2 + 6z_1z_2 + 12z_2{}^2) \qquad (9)$$

where

$$z_1 \equiv 1 - \frac{1}{n}\sum_{i=1}^{n} ln\frac{x_{(i)}}{x_{(n)}}$$

and

$$z_2 \equiv \frac{1}{2} - \frac{1}{n}\sum_{i=1}^{n}\frac{x_n}{x_i}$$

which is asymptotically distributed under the null as a chi-square with 2 degrees of freedom. An appealing feature of the test is that it is locally optimal if the alternatives are other power laws. Table 3.1 below presents the significance points for LMZ value.

Table 3.1 Significance points for LMZ[a]

n	10	15	20	25	30	50	100	200	∞
Level									
5%	6.19	6.14	6.09	6.08	6.03	5.98	5.98	5.99	5.99
10%	4.38	4.41	4.43	4.45	4.46	4.49	4.56	4.58	4.61

[a]Source: Urzua (2000) Monte Carlo simulations using the inversion method, and after 100,000 replications.

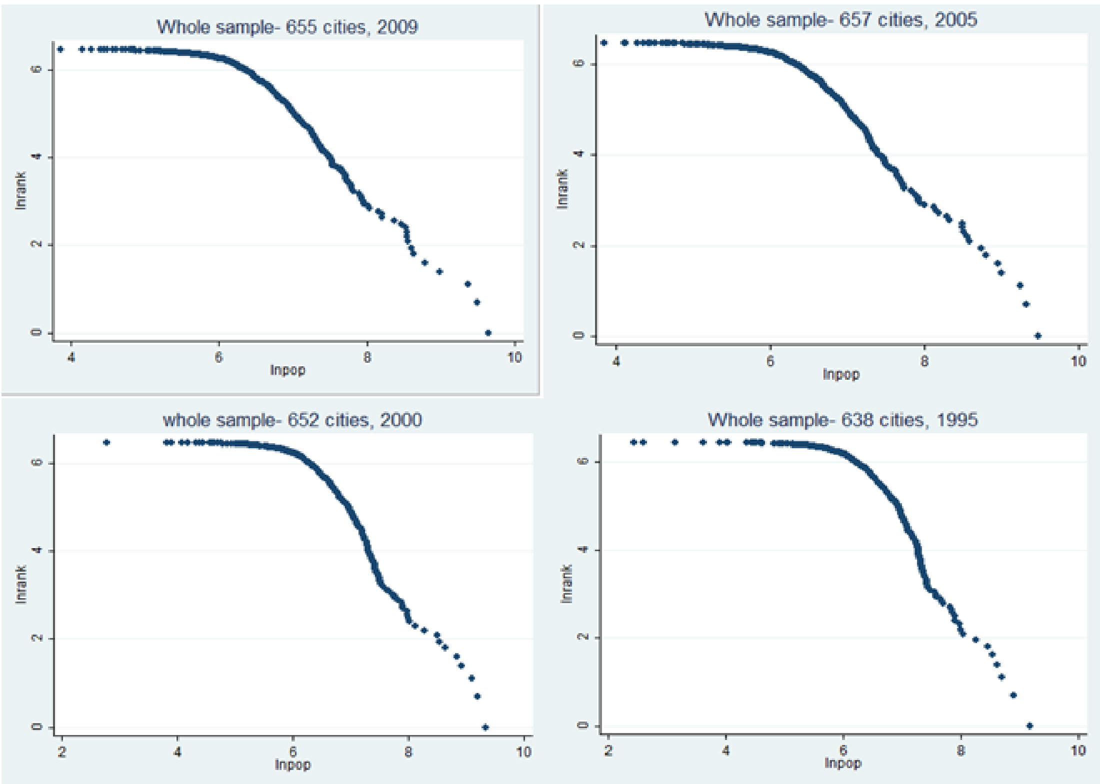
In the end, we test for the exponent not only for the whole sample, but also firstly for the prefecture-level cities and county-level cities. Secondly for the truncated samples that rank Top50, Top100 and Top 280 subsamples to see the Zipf exponent, as many literature find that Zipf's law at least holds for upper truncated sample. In addition, following Eeckhout 2004 and Anderson and Ge 2005 we regress the logarithm of rank on logarithm of size using the panel dataset to compare with the repeated cross sectional results.

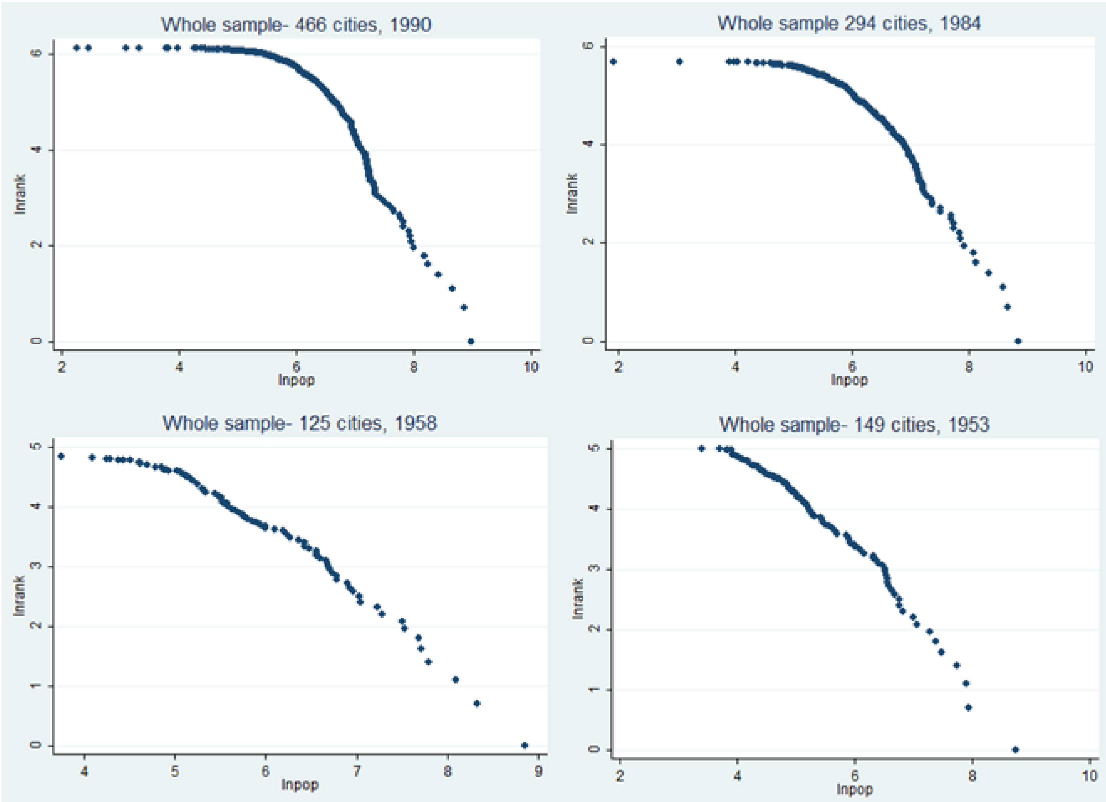
3.4 Empirical Results

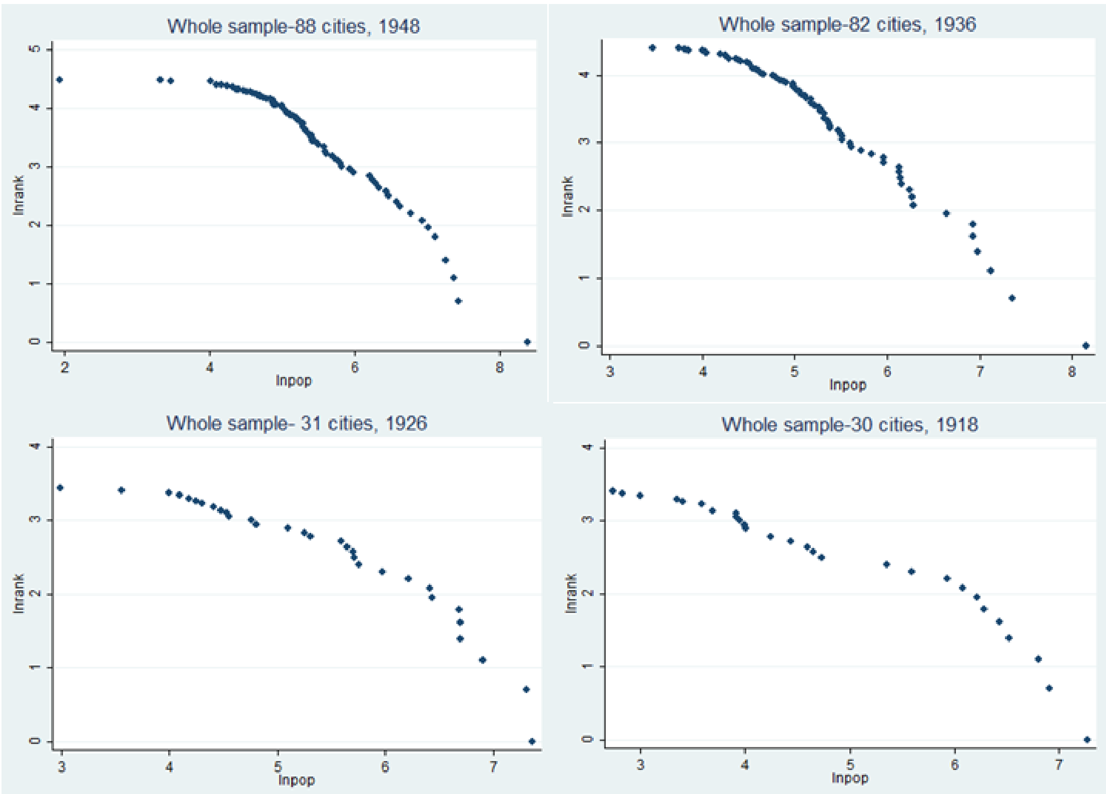
3.4.1 Zipf's Plots

With respect to results, first of all, we will see the conventional log-log Zipf's plots of rank and size to have some intuitive grasp of Zipf's law in Figure 3.2. If Zipf's law holds, we expect to find a nearly straight line between rank and size (with the slope equals to -1). Figure 3.2 below shows the evolution of Zipf's plot over time from 1879 to 2009, using a set of graphs that scatter the logarithm of city rank (ranking by population within a year) against the logarithm of city size (population). Overall, at the first glance, none of the graph indicates Zipf's law. For the relatively small cities, the city size distribution is much flatter than Zipf's law predicts, indicating that small cities are distributed much less evenly as Zipf's law statement, i.e. the discrepancy of size difference is quite large in small cities.

Figure 3.2 shows the Chinese city size distributions from 1879 to 2009 for the whole sample (all the cities available). One can notice that from 1879 (end of the 'Qing' Dynasty) to 1953 (the beginning of PRC China), despite the WWI, WWII and Civil War shocks the city size distribution surprisingly shows a tendency to Zipf's law. In other words, during this period, Zipf's plots are more and more closing to a straight line; especially in 1953, the line is close to a linear line with a slope -0.89 (shows in next section 4.2 Table 3.4- Pareto exponents). Chinese city size distribution in this period (before PRC China and back to the end of 'Qing' Dynasty) has not been tested in previous literature, but the result is surprisingly close to Zipf's law, in spite of the various war shocks, which is consistent with the relative studies explaining Zipf's law that if the shock is not permanent to the growth then in the long run in steady state Zipf's law emerges (Gabaix, 1999). From 1958 to 2009, this is a more stable period of 'new' China, obviously Zipf's law does not hold for the whole sample in this period because of the emergence of small cities showing as a quite flat part in the lower tail of Zipf's plot in every graph after 1958. However, we can notice that the middle part of Zipf's plot seems to follow a straight line in every graph after 1958 (incl.), which means except for the new entry quite small cities and the quite large cities, the medium-sized cities seem to follow Zipf's law.




Whole sample- 466 cities, 1990
Whole sample 294 cities, 1984
Whole sample- 125 cities, 1958
Whole sample- 149 cities, 1953
lnrank
lnpop


Whole sample-88 cities, 1948
Whole sample-82 cities, 1936
Whole sample- 31 cities, 1926
Whole sample-30 cities, 1918
lnrank
lnpop

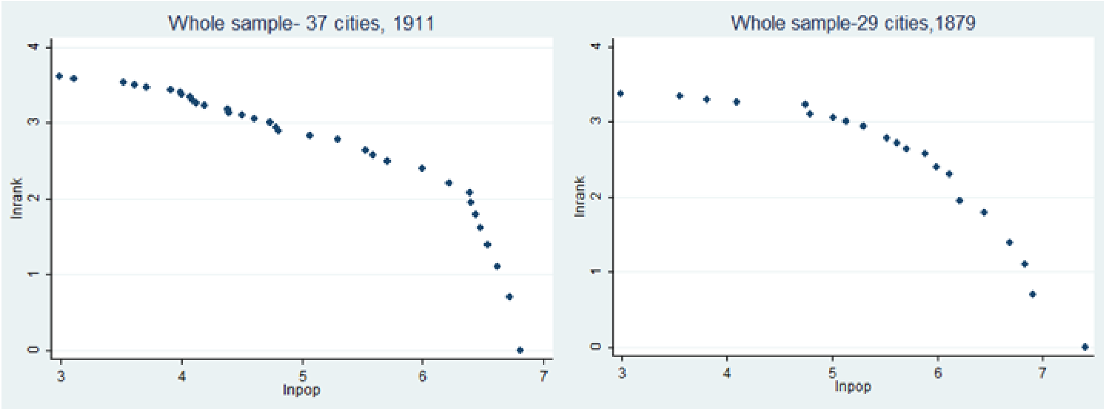


Figure 3.2: Zipf's plot for the whole sample over time.

Figure 3.3 below plots the logarithm of city rank and logarithm of city size for the 'middle-ranked' cities[29] that contribute to form the straight-line part in Zipf plots over time. We still analyse chronologically. Firstly, from the whole sample plots in Figure 3.2 we notice that in the Zipf plot graph for year 1879 and 1911 (war chaos in the end of 'Qing' Dynasty and WWI period), each of the Zipf plot approximately consists of two straight lines with different slopes for the lower tail and upper tail parts respectively. Therefore we are not showing the 'middle-ranked' cities for these two years, instead we show the Zipf plots for the high ranked (upper tail) and the low ranked cities (lower tail) respectively in Figure 3.3- Panel A. For these two years, Zipf plots for large cities also show a straight line (the left part of graph for these two years). It seems like Zipf's law but the absolute value of the slope is much smaller than 1 (0.79, shown in Table 3.2 Pareto exponent in section 4.2), meaning that city size distribution for large cities (upper tail) is not as even as Zipf's prediction, i.e. large cities are too large and small cities are too small. Similar for the lower tail -

[29] In fact we are not defining the 'middle-ranked' cities; actually we find the roughly rank span which consists of the straight-line part in Zipf plots, showing in the title of each graph in Figure 3.

medium and relatively small cities- city size distribution also follows a straight line with slope still much smaller than 1 (0.74), again indicating that the rest of the cities are distributed less evenly than Zipf's law prediction.

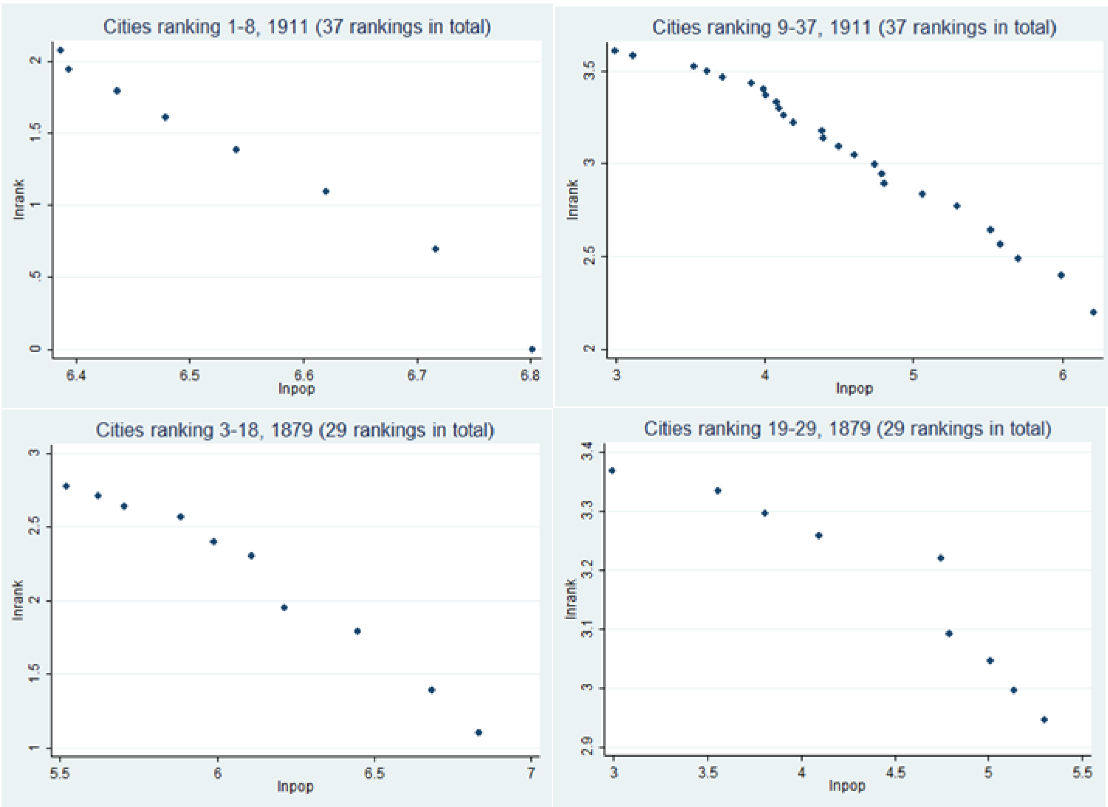


Figure 3.3_ Panel A: Zipf's plot for city hierarchy closing to straight line over time.

Secondly, from 1918 to 1953 (the graph below, Figure 3.3- Panel B), except for several extremely large cities (ranking from 1 to around 10) and a few extremely small cities (inverse ranking 1 to round 10), almost all of the cities in the distribution follow a log-log straight line with the slope smaller than 1 (absolute slope), but the slope is greater than 1879 and 1911 graphs (the average absolute Pareto exponent is 0.87, show in Table 3.2, section 4.2). This shows that cities' size distribution from

1918 to 1953 (WWII and Civil War period) is still less even than Zipf's law predicts, but more even and closer to Zipf's law than the 1879 to 1911 period.

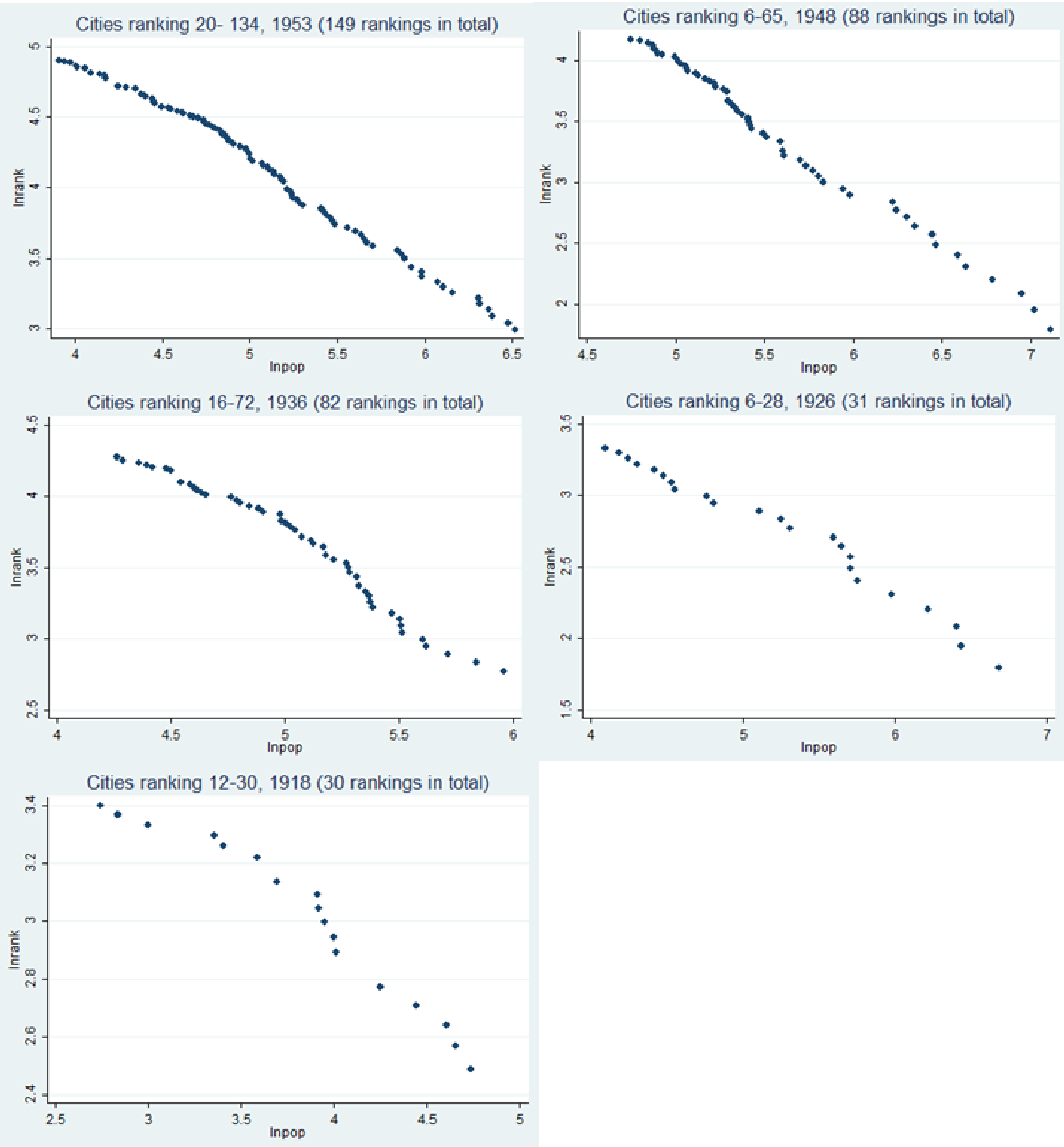


Figure 3.3_ Panel B: Zipf's plot for city hierarchy closing to straight line over time.

Thirdly, from 1958 to 2009 (graph below, Figure 3.3- Panel C), the 'new' China's stable development period, almost all of the medium-sized cities ranking from around the top 20% to the top 70% follow a log-log straight line, which indicates that there is a linear inverse relationship between rank and size. But the absolute value of the slope is still a bit smaller than 1 in every graph, but it does greater than the period before 1953 in Figure 3.3-Panel A and B (far less even) and increase from 1958 to 2009 to around 0.8, which indicates that medium-sized cities are distributed more and more evenly as Zipf's predicts (slope should be -1). These results are all consistent with the economic and urban development situation in these periods (after the establishment of PRC China, there is a relatively stable economic and political environment for cities to grow) and also in line with the Zipf's law evolution (city size growth process needs time to converge to Zipf's law, Gabaix 1999).

To be noted that, for the upper tail i.e. large cities, contrary to other studies of developed countries that Zipf's law holds at least for the large cities (Krugman, 1996; Gabaix ,1999a; Dobkins and Ioannides, 2000; Ioannides and Overman, 2003; Black and Henderson, 2003 and Eeckhout, 2004 for developed countries studies; Song and Zhang, 2002 and Gan, Li and Song, 2006 for Chinese case), in our study Zipf's law does not hold in upper tail during the whole period. The Pareto exponents for large cities are consistently greater than 1 as the next section shows.

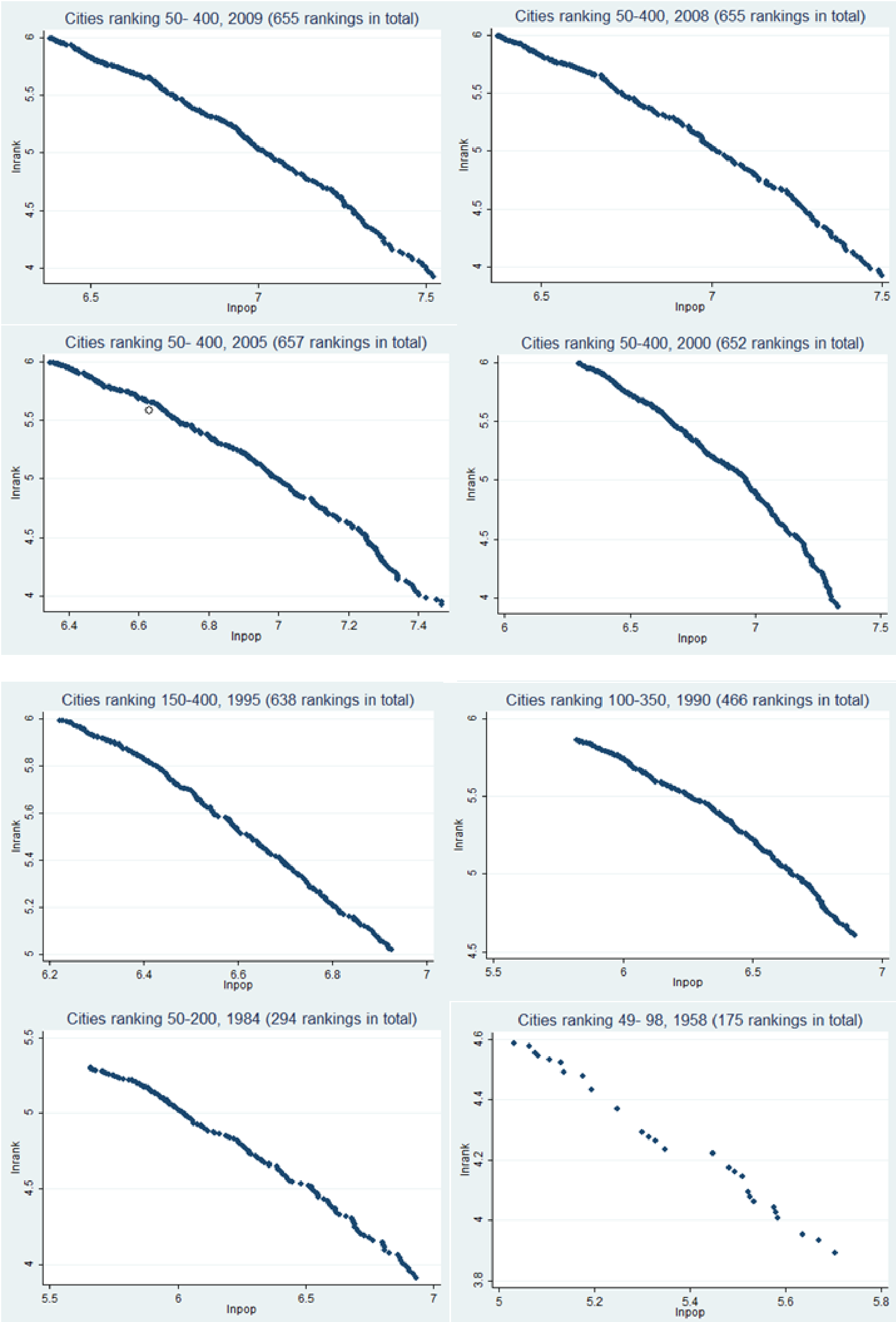


Figure 3.3_ Panel C Zipf's plot for city hierarchy closing to straight line over time.

3.4.2 Repeated Cross-Sectional OLS Regression and Lagrange Multiplier (LM) for Pareto Exponent

(1) Testing for the whole sample

The other way to test for Zipf's law usually is the regression method, regress the logarithm of city rank against the logarithm of city size and test for whether the Pareto exponent equals to -1. It has been customary to test for the Pareto exponent using OLS estimation, although some studies criticize that it is inefficient and biased in testing for city ranks (as mentioned in section 3). We firstly show the estimated OLS Pareto exponents and then a new method employed: Lagrange multiplier (LM) test (Urzua, 2000 and 2011).

Firstly, Table 3.2 below shows the repeated cross sectional regression results for estimated Pareto exponents and LMZ test statistics for the whole sample, prefecture-level cities and county-level cities respectively. Columns OLS(1) and OLS(2) indicate the regression of Eq.(4) and Eq.(7) , the conventional OLS and corrected OLS regression. LMZ statistic values are also reported and the significance critical value is in Table 3.1. We can see the evolution of Zipf's estimated coefficient over time in Figure 3.4 below.

Table 3.2 Evolution of Pareto exponent

	Whole sample				Prefecture-level cities			County-level cities		
Year	No. of cities	OLS(1)	OLS(2)	LMZ	No. of cities	OLS(1)	OLS(2)	No. of cities	OSL(1)	OLS(2)
1879	29	-0.717	-0.792	151.254	28	-0.717	-0.792			
1911	37	-0.687	-0.743	154.880	34	-0.703	-0.762	3	-0.346	-0.352
1918	30	-0.601	-0.663	105.669	27	-0.613	-0.679	3	-0.432	-0.443
1926	31	-0.690	-0.759	155.996	29	-0.689	-0.758	2	-0.413	-0.423
1936	82	-1.006	-1.069	397.756	80	-1.005	-1.068			
1948	88	-0.862	-0.912	468.276	86	-0.863	-0.913	2	-0.402	-0.405
1953	149	-0.890	-0.926	633.193	128	-0.909	-0.948	21	-0.492	-0.494
1958	125	-0.954	-0.997	664.544	114	-0.984	-1.031	11	-0.474	-0.476
1983	192	-1.043	-1.077	1007.753	170	-1.070	-1.106	22	-0.662	-0.666
1984	294	-0.970	-0.992	1245.110	227	-1.203	-1.235	67	-0.420	-0.421
1985	324	-0.918	-0.937	1366.795	237	-1.246	-1.278	87	-0.414	-0.416
1986	320	-0.955	-0.976	1383.598	231	-1.261	-1.294	89	-0.437	-0.438
1987	382	-0.953	-0.971	1642.794	240	-1.319	-1.351	142	-0.552	-0.554
1988	433	-0.983	-1.000	1870.302	252	-1.386	-1.418	181	-0.593	-0.595
1989	447	-1.001	-1.019	1966.603	250	-1.418	-1.451	197	-0.608	-0.610
1990	466	-1.016	-1.033	2086.941	253	-1.441	-1.474	213	-0.629	-0.631

1991	478	-1.030	-1.047	2154.093	254	-1.461	-1.494		224	-0.647	-0.649
1992	516	-1.047	-1.063	2338.841	257	-1.497	-1.530		259	-0.691	-0.693
1993	570	-1.078	-1.094	2601.236	261	-1.539	-1.572		309	-0.732	-0.734
1994	620	-1.097	-1.112	2843.818	262	-1.587	-1.621		358	-0.750	-0.753
1995	638	-1.116	-1.131	2940.598	260	-1.624	-1.658		378	-0.765	-0.767
1996	664	-1.140	-1.155	3110.162	261	-1.632	-1.665		403	-0.792	-0.794
1999	663	-1.090	-1.105	3459.414	262	-1.245	-1.269		401	-0.874	-0.876
2000	652	-1.176	-1.192	3354.237	259	-1.517	-1.549		393	-0.883	-0.886
2001	662	-1.116	-1.131	3513.878	268	-1.331	-1.358		394	-0.859	-0.862
2002	655	-1.130	-1.145	3509.038	274	-1.387	-1.415		381	-0.822	-0.824
2003	654	-1.143	-1.159	3654.414	277	-1.417	-1.445		377	-0.811	-0.813
2004	653	-1.163	-1.179	3984.158	279	-1.435	-1.463		374	-0.823	-0.825
2005	657	-1.176	-1.192	4257.821	278	-1.524	-1.555		379	-0.803	-0.805
2006	659	-1.183	-1.200	4290.257	279	-1.536	-1.567		380	-0.810	-0.812
2007	653	-1.184	-1.201	4234.632	279	-1.534	-1.566		374	-0.806	-0.808
2008	654	-1.188	-1.205	4244.710	280	-1.520	-1.552		374	-0.810	-0.812
2009	653	-1.182	-1.198	4235.137	280	-1.523	-1.554		373	-0.798	-0.800

Notes:

For OLS regressions, Newey-West standard errors controlling for heteroskedasticity reported.

OLS (2) is the correction of OLS following Gabaix and Ibragimov (2011) to regress equation (7).

Standard errors, Cook-Weisberg test for heteroskedasticity and Ramsey RESET test for omitted variables see Appendix Table A2.

Test for alpha=-1 which is the basic test for Zipf's law and the joint test which is the Alperovich joint test that alpha=1 and Constant= largest city in the basic Zipf regression, also see Appendix Table A3.2.

For OLS(1) and OLS(2) the "*" means significantly different from -1.

Significance points for *LMZ* : 5.98 at 5% level, 4.49 at 10% level.

To be noted that it is unlike normal estimation results testing for the estimated coefficients whether significantly different from 0, in contrast, we are interested in whether the estimated Zipf's exponent significantly different from 1. Results show that every single estimated Zipf's exponent is significantly different from 0 at the 1% level, where in normal result table, there should be a significance indicator "***". However, we do not show this significance indicator because that is not of our interest.

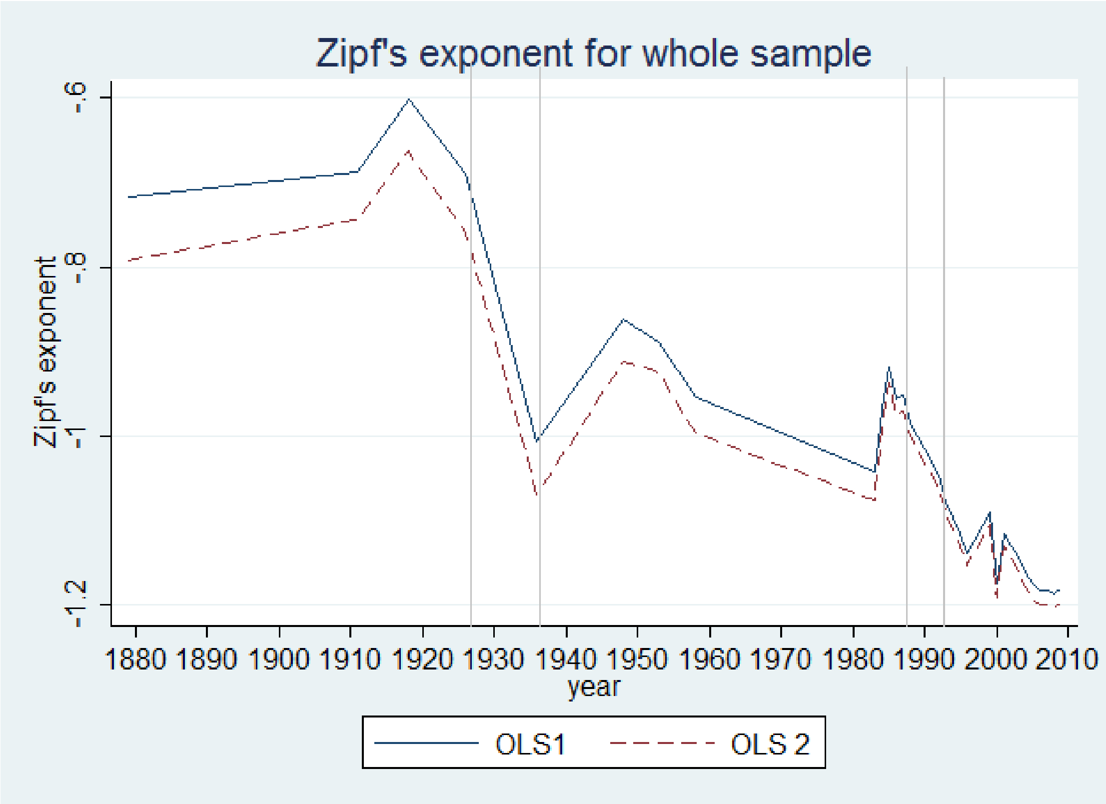


Figure 3.4_Panel A: Evolution of Zipf's exponent over time for the whole sample

For the whole sample, we can see from Figure3.4- Panel A (above) and column 3 in Table3.2 (OLS(1)), from 1879 to 1926, the absolute value of Zipf exponents are less than 1 (around -0.6 and -0.8) indicating the cities are distributed less evenly than Zipf's law predicts, i.e. in terms of size, there is a relatively large disparity between large cities and small cities, which is also consistent with the Zipf's plot mentioned before. This large disparity between large and small cities may be caused by the chaotic multi-governments, weak economy and wars during this period. Urban residents tend to concentrate in large cities to take the advantage of 'foreign settlement' in large cities to avoid the war, and to seek better economic and quality of life environment. This might make more differences between large and small city sizes.

Then from 1926 to 1936 the absolute value of the Zipf's exponent increases from 0.8 to 1, as the relatively united National Government established in 1926 and cities begin to grow more even. Until 1987 the Zipf exponent is between -0.86 and -0.96 (except for two years reach -1.01 (1936) and -1.04 (1983)) we can see that the distribution of cities starts growing more and more even and even for some specific years the strict Zipf's law holds (1936, 1983), due to the more and more stable economy and urban development.

Consistent with the previous literature explaining Zipf's law (Gabaix, 1999), after some time for urban growth, Zipf's law is found between 1988 and 1992, with Pareto exponent from -0.98 to -1.047.

Then from 1993 to 2009 the absolute value of Zipf exponent is relatively stable and consistently greater than 1, increasing over time from -1.078 to -1.19, showing that Chinese cities are growing more and more even in terms of size due to the stable economic and urban growth, which is also consistent with Peng (2010) who tests from 1999 to 2004.

On the whole, the absolute value of Zipf exponents increases over time from 1879 to 2009 (-0.6 to -1.19) with some fluctuation, i.e. the slope of Zipf plot is increasing over time, which indicates that the city distribution is evolving from less even to more even, i.e. differences in city sizes are decreasing. This is consistent with the economic development of China during these years and also indicates that over these 130 years, from the end of the 'Qing' Dynasty to the present day, city growth is not only concentrated in large cities. Indeed, medium and small cities might grow faster during some periods, perhaps when the size of large cities reach some certain level, as suggested by Sequential City Growth theory in a later chapter.

To conclude for the whole sample, the Zipf's coefficient generally increases from 1879 to 2009 (ranging from 0.6 to 1.27) and there is a linear relationship between rank and size for almost all of the medium-sized cities. Especially, Zipf's law is found to hold for the end of 1980s and early 1990s (Pareto exponent is around -1, from -0.98 to -1.03). These results are also consistent with previous studies: Song and Zhang (2002) study Chinese cities' Pareto coefficient for 1991 and 1998; Gan, Li and Song (2006) study for 1985 and 1999 and they've found Zipf coefficient is very close to -1, in our case for these years the coefficient is close to -1 as well. The difference from other studies is that they haven't checked the Pareto exponent consistently for every year (normally just 2 years. 8 years at the most in the case of Anderson and Ge, 2005). However, we need to know the evolution of the Pareto exponent over time to study

Zipf’s law and city size distribution, since either Zipf’s law or city size distribution may need time to converge.

We find that before 1989 the absolute value of the Zipf’s coefficient is often smaller than 1, while from 1989 afterwards it is generally greater than 1, indicating that before 1989 the whole Chinese city size is not as even as Zipf’s prediction but might have a trend to develop more even and then cities grow as evenly as Zipf’s law prediction, but finally they develop a bit more evenly as Zipf’s law states.

To be noted that from the mid-1980s to the end of 1980s, there is a decrease of Zipf’s exponent (absolute value) for the whole sample, this may because of the first stage of ‘Economic Reform 1979’ which encourage some specific regions or cities to grow first (they may receive more favourable development policies from central government than other cities), and then the adjacent regions or cities expect to develop by the leading of early developed regions or cities (for instance, the designed ‘special economic zone’, etc.). This first stage is reflected from this period, and the policy prior supporting for some specific cities leads to the increasing disparity between large and small cities, thus decrease the Zipf’s exponent. After the first stage of ‘Economic Reform’, from early 1990s, the second stage is reflected. When other cities start to develop, the differences between cities decrease, thus, Zipf’s exponent increases.

(2) Testing for Prefecture-level cities and county-level cities

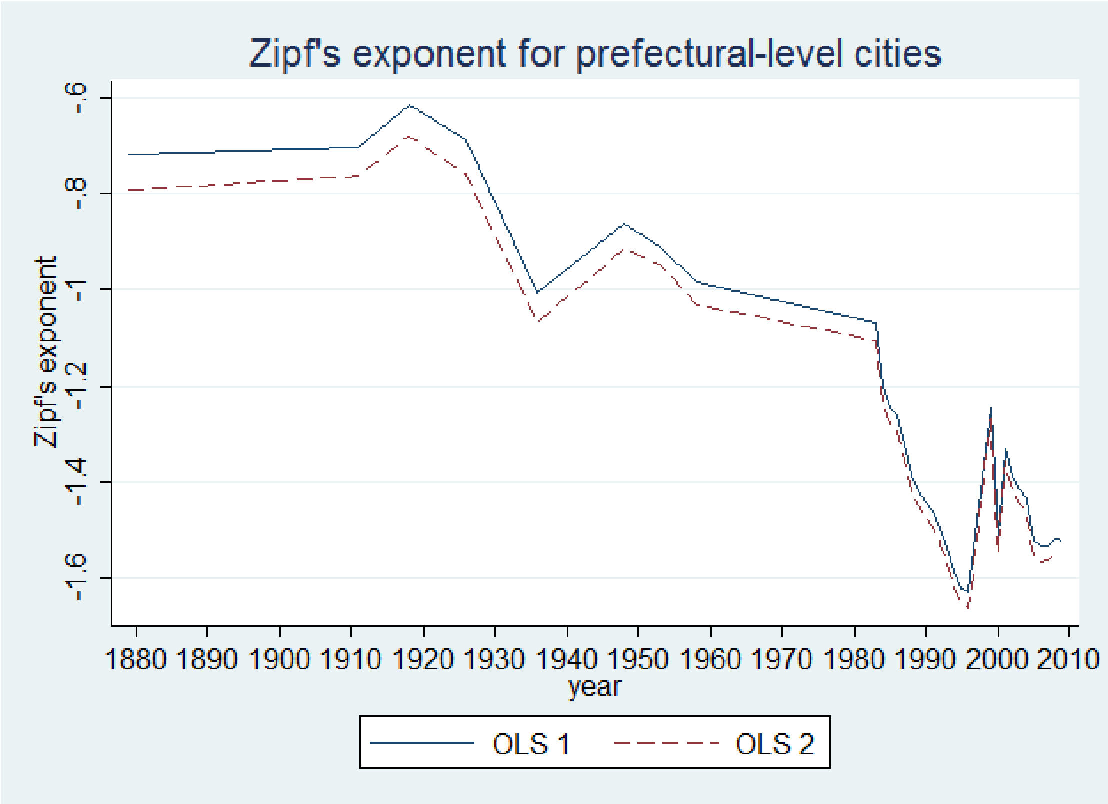


Figure 3.4_Panel B: Evolution of Zipf's exponent over time for prefecture-level cities

Next we split the whole sample in to prefecture-level cities and county-level cities according to the national administrative bureau, as these two kinds of cities are quite different in terms of history, size and policy direction. Firstly, with respect to prefecture-level cities, they are designated as prefecture-level cities by the government for their relatively long history of economic and human agglomeration. Besides, prefecture-level cities have received relatively more favorable policies in terms of either economic or urban development. Therefore we might expect the distribution of prefecture-level cities' to be more even and stable than the distribution of other cities. Figure3.4- Panel B and Table 3.2 column 7 above both show the evolution of the Pareto exponent for prefecture-level cities. As we expected, the city size distribution shows a typical evolution of Zip's law:

(1) Less even: at early war chaos stage, before 1953, the prefecture-level city size distribution is less even as Zipf's law predicts, as the absolute value of the Pareto exponent is firstly around 0.6 to 0.7 before 1926 (incl.)- WWI period-, then grows to around 0.8 to 0.9 in the second world war period- WWII and the Civil War- from 1936 to 1953. This may because of the unbalanced urban growth during undeveloped period.

(2) Zipf's predicts: from 1958 to 1983, 'new China' (People's Republic of China) provides a relatively stable political environment. But from 1958, the 'Hukou' system was launched which restricts migration from rural areas to the cities. However, Zipf's law emerges in this period, with Pareto exponent around -0.98 to -1.07, i.e., cities are distributed evenly as Zipf's law predicts; the log rank and log population plot would show a straight line with slope -1. The possible explanation of the validity of Zipf's law under migration restrains is that what the 'Hukou' system strictly restrict is the migration from rural to urban area (details about 'Hukou' system was described above in section 2.2.3), not the intra-city migration. This

policy seems do not affect the result, but is still addressed here because this is one of the uniqueness of analysing Chinese cities.

(3) More even: from 1980, the 'Economic Reform' plays a significant role not only in the national economy but also in the urban growth process. It generates more even urban development for prefectural level cities because the promotion of economic growth for all cities increases the Zipf's exponent and thus produces more even city size distribution. Specifically, 'Economic Reform' enhance the industrialisation and openness level of cities, especially for medium and small cities for instance, Shenzhen (Guangdong) developed from a small city to a relatively large city after 'Economic Reform' as it was one of the designed 'special economic zone'. As the development of economy promoting by 'Economic Reform', prefecture-level cities grow more evenly than Zipf's law predicts (the slope in Zipf's plot is increasing), which is reflected in the increasing Pareto exponents. This may because that the level of income increase, thus the differences between cities decrease.

Specifically, from 1984 to 1996, the Pareto exponent increases constantly from -1.2 to -1.63, then shows fluctuation between -1.2 and -1.5 in the end of 1990s to the start of 2000s, finally it is relatively stable at -1.52 or -1.53 from 2005 to 2009. To conclude, the evolution of city size distribution in prefecture-level cities becomes more even over time, until recent years it shows a comparatively stable city size distribution. Zipf's law emerged in this process.

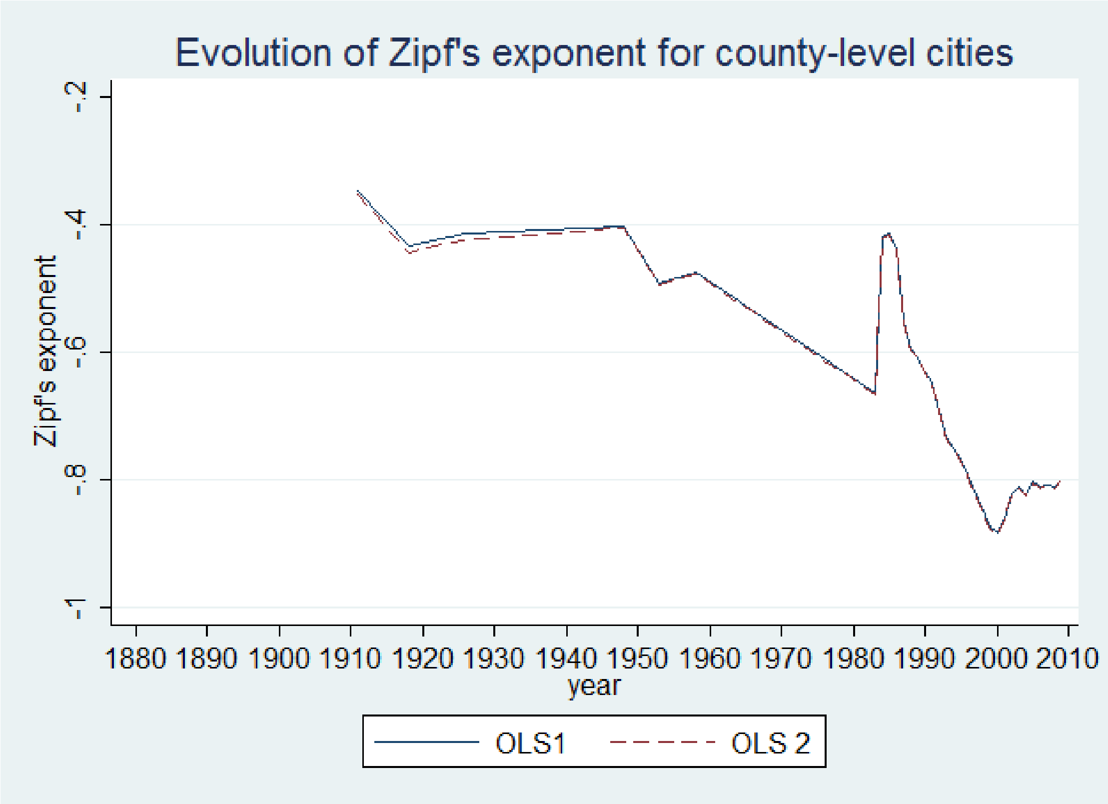


Figure 3.4_Panel C: Evolution of Zipf's exponent over time for county-level cities

On the contrary, for county-level cities, in Figure3.4 PanelC above and column 10 in Table 3.4 below, the absolute value of the Pareto exponent is consistently significantly smaller than 1 showing that a large city size disparity exists in county-level cities.

Specifically, (1) before 1986 the Pareto exponent is always around -0.4 (two exceptions for -0.34 and -0.66), which means that the size differential between large and small cities is quite big, i.e. city size distribution is quite unequal. This is partly because the number of county-level cities was few and urban growth policy did not favour them during that period, large cities got the priority for city growth.

(2) 1987 to 2000, the effect of 'Economic Reform' started to reveal itself; it also pushed the urban growth of county-level cities. The Pareto exponent rapidly increases within two decades, from -0.55 (1987) to -0.88 (2000) which is much closer to the Zipf's law prediction of -1. County-level cities in the last two decades of the 20^{th} century are developed much more evenly, i.e. the size disparity between large and small cities is decreasing, due to the 'Economic Reform' (generates more balanced development of cities, decrease the differences between city sizes by increasing the income level of over all cities) and accordingly greater migration between cities and between rural and urban areas.

(3) After the year 2000, county-level city size distribution is relatively stable, with the Pareto exponent relatively stable around -0.79 to -0.85. To conclude, county-level cities are distributed quite unequally before 'Economic Reform', but become more even within the 1980s and 1990s. After entering the 21^{st} century they show relative stable distribution with not too large a disparity between the lower and upper tails.

To be noted that there is also a decrease of Zipf's exponent in the mid-1980s to the end of 1980s, i.e. the difference between city sizes increase. This may also because of the first stage of 'Economic Reform' where it supports some specific cities to grow first, thus increase the disparity between city sizes.

In addition, Figure 3.5 below shows the size distribution geographically for prefecture-level cities and county-level cities in 2009. From Table 3.2, we know that for 2009 the Pareto exponent is -1.55 and -0.8 for prefecture-level cities and county-level cities respectively, which means prefecture-level cities are distributed a bit more equally than Zipf's prediction while county-level cities are distributed a bit less evenly than Zipf's prediction.

Within prefecture-level cities, the first ranked city has roughly $\sqrt[1.55]{2} = 1.56$ times the population of the second ranked city and has roughly $\sqrt[1.55]{3} = 2.02$ times the population of the third ranked city, etc. Within county-level cities, the first ranked city has $\sqrt[0.8]{2} = 2.38$ times population of the second ranked city and has $\sqrt[0.8]{3} = 3.95$ times population of the third ranked city, etc. The disparity between large and small cities is relatively small in prefecture-level cities' sample but large in county-level cities' sample. We can also see this from the graphs below. For prefecture-level cities, the transition from the biggest circle to the smallest circle is gradual, while for county-level cities the transition is less gradual; in other words, for prefecture-level cities there are big, medium big, medium small and small circles, but for county-level cities, circles are concentrated in either big or small. Geographically, prefecture-level cities are spread more evenly than county-level cities.

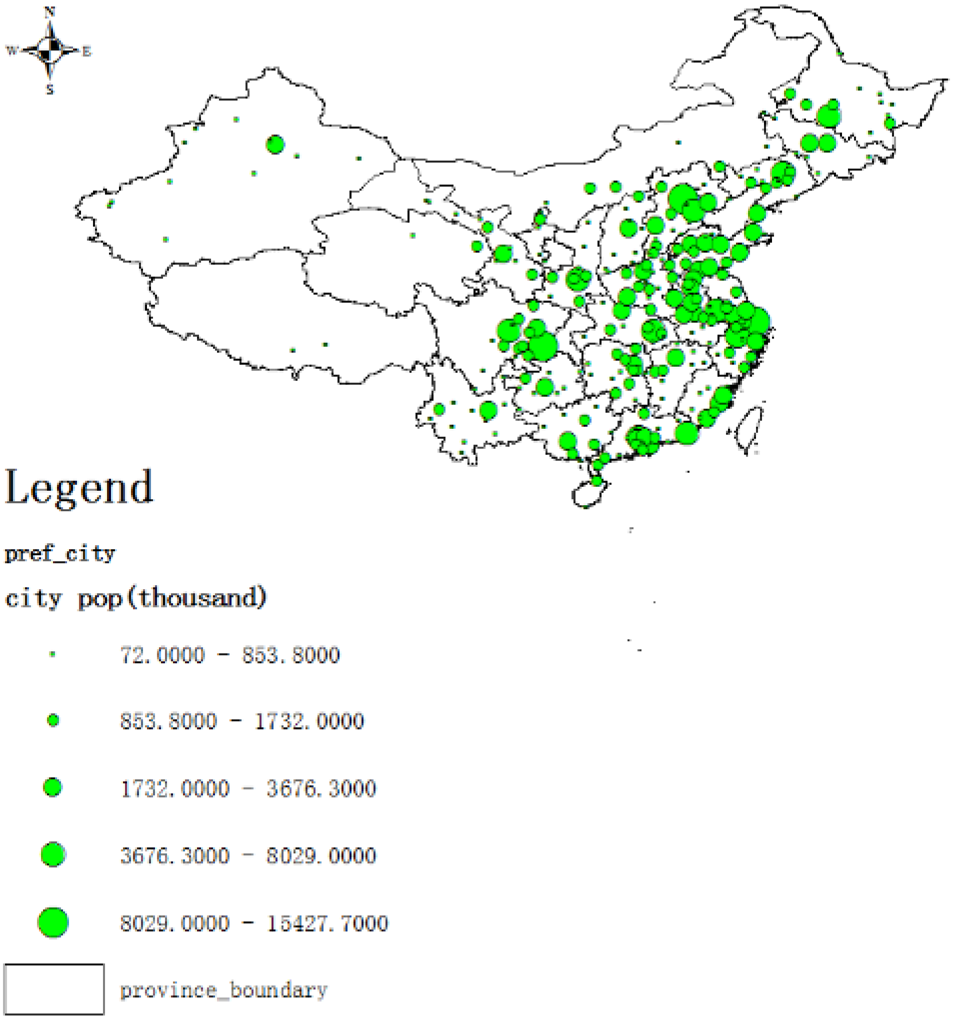


Figure 3.5_Panel A: Prefecture-level city size distribution geographically.

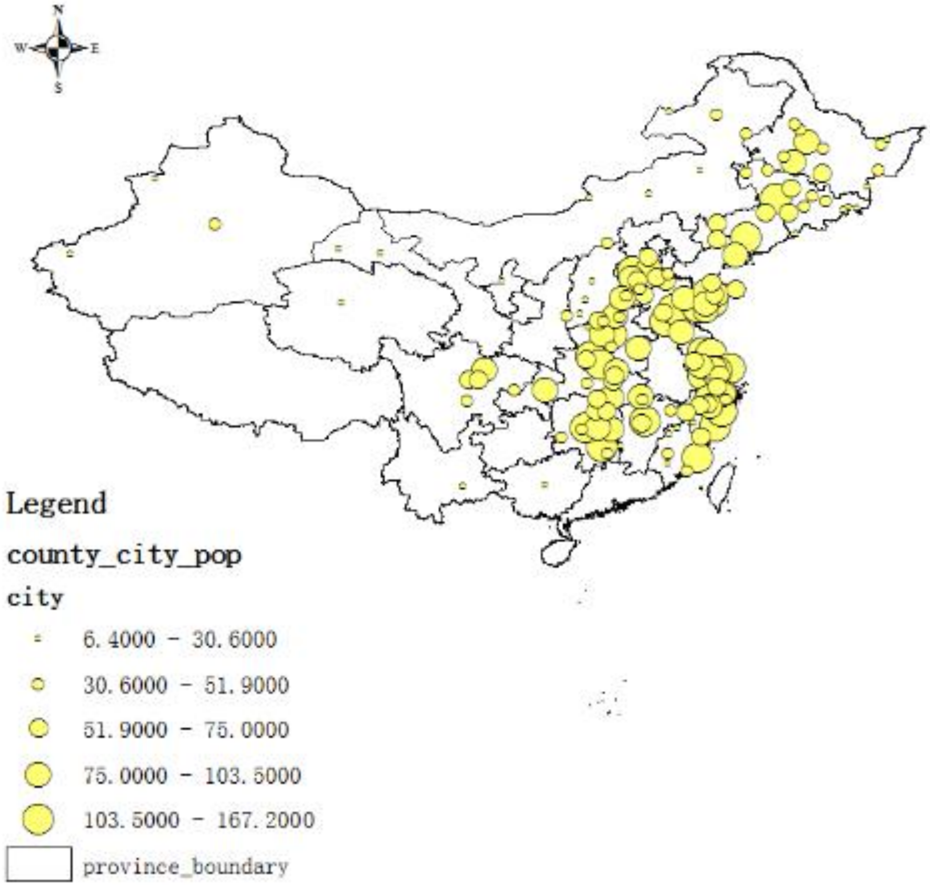


Figure 3.5_Panel B County-level city size distribution geographically.

(3) Testing for truncated samples

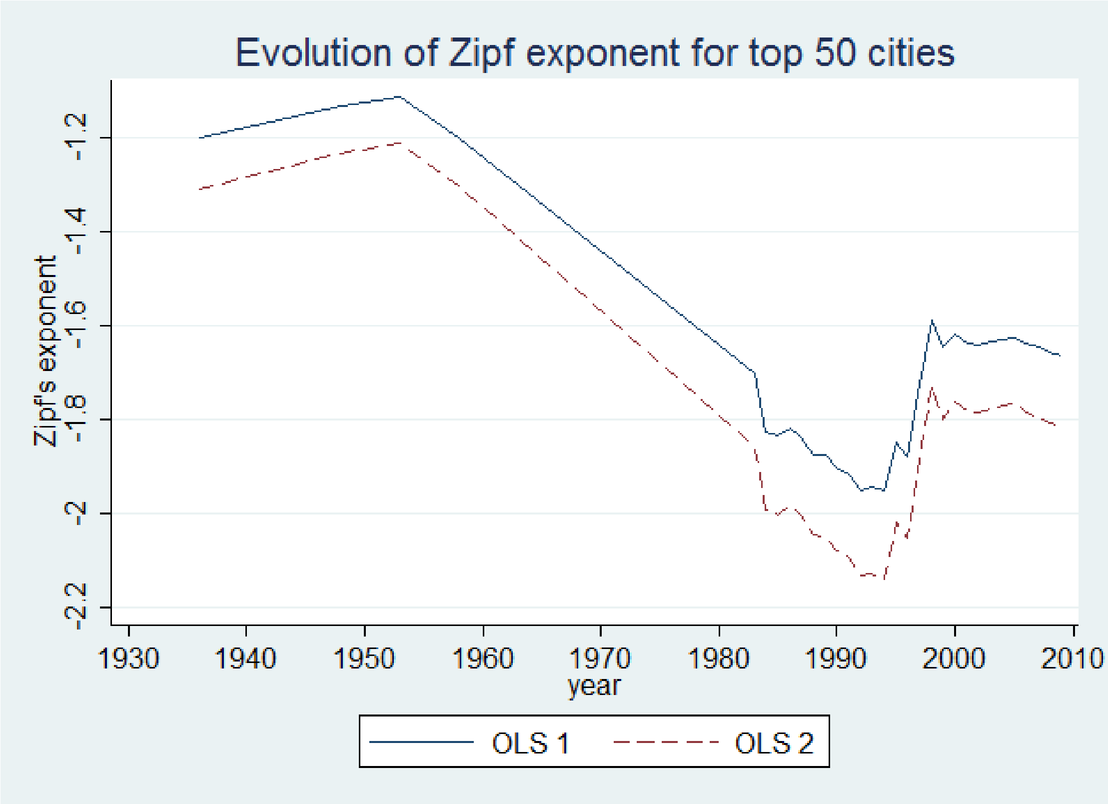


Figure 3.6_Panel A: Evolution of Zipf's exponent for truncated samples (Top 50 cities)

The previous literature has mentioned that the Pareto exponent is quite sensitive to the truncation point (Rosen and Resnick, 1980; Black and Henderson, 2003; Eeckhout, 2004. Song and Zhnag, 2002; Peng, 2010 for Chinese case). We also test for the truncated sample by ranking the top 50 cities, ranking the top 100 cities and ranking the top 280 cities. In Table 3.5 below, we run regressions across three subsamples to explore city size distribution in truncated samples and also we ensure the number of cities in each year is fixed. Besides, Figure 3.6 above illustrates the evolution of the estimated Pareto exponent over time.

Firstly, consider the top 50 sample, we find that the top 50 cities are often distributed more evenly than Zipf's prediction, especially after the 'Economic Reform'. More details in Figure 3.6- Panel A and column 2 in Table 3.3, we notice:

(1). 1936 to 1958, during the WWII and Civil War, the top 50 large cities are distributed reasonably evenly (around -1.13), not too much disparity between large and small cities.

(2). from 1983 to mid of 1994, absolute value of the slope of this linear relationship is increasing a bit within this decade from -1.7 to -1.95. The top large 50 cities are growing much more equally than Zipf's law predicts, i.e. the disparities between large and small cities decrease again. This might be because the urban growth of these top large cities are affected equally by the economic development policy due to the 'Economc Reform', where the 'Economic Reform' enhance the income level of all cities thus reduce the differences between city sizes.

(3). at the late stage of the 1990s, the Pareto exponent decreases from -1.84 to -1.59, which means that top 50 cities become less even, some cities may grow much faster than others. (4). after 2000, the city size distribution of the top 50 large cities is relatively stable and a bit more equal than Zipf's law predicts, reflected in the Pareto exponent staying quite stable around -1.65. This result is suprisingly consistent with the county-level city's sample (relatively small cities), but one can tell from the Zipf's plots for the years after 2000, the slope of upper and lower tails are quite stable. To conclude, Zipf's law is not found in top ranking 50 cites since over time they are distributed much more equally than Zipf's prediction.

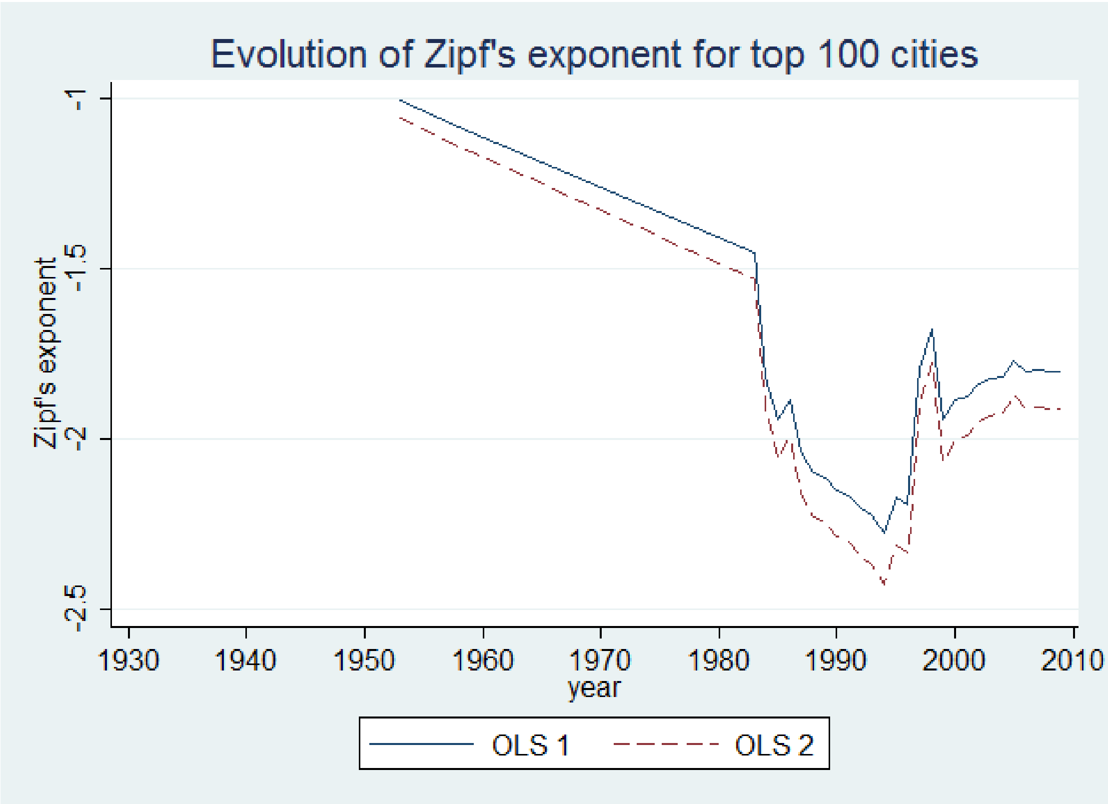


Figure 3.6_Panel B Evolution of Zipf's exponent for truncated samples (Top 100 cities)

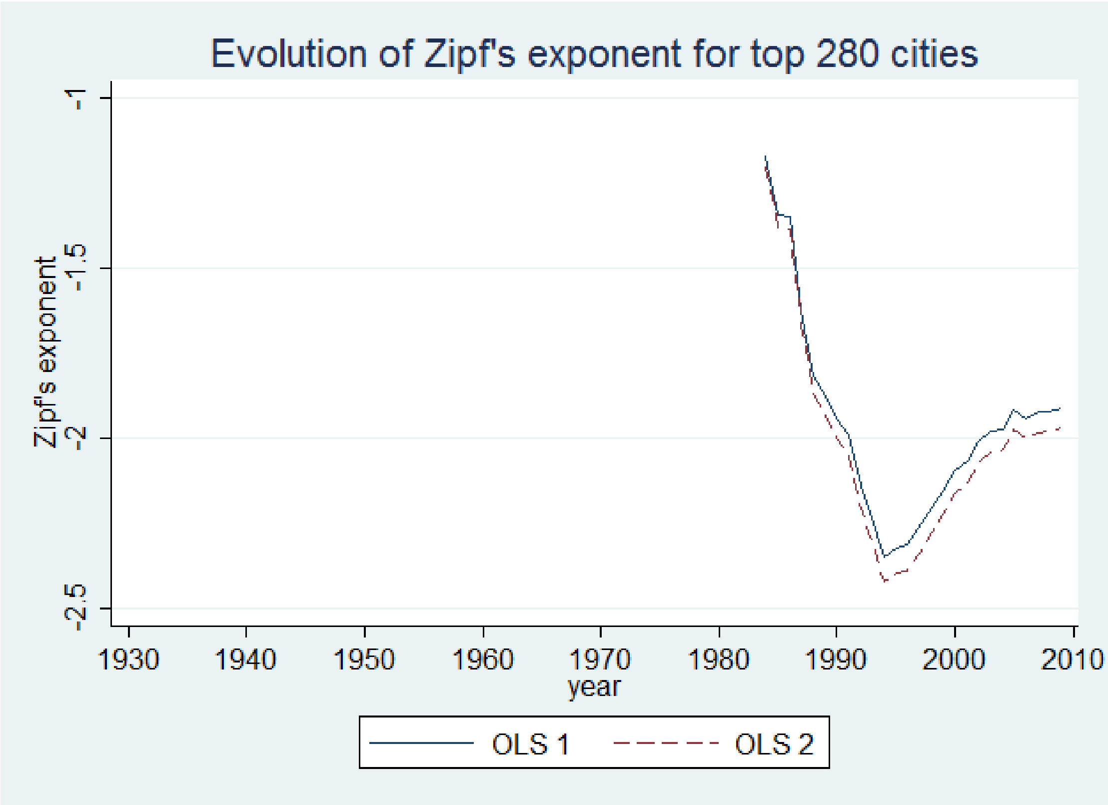


Figure 3.6_Panel C: Evolution of Zipf's exponent for truncated samples (Top 280 cities)

Secondly, for the top ranking 100 cities' subsample, we see quite similar distribution and shape with the top 50 cities' sample from 1953 to 2009 by comparing Figure 3.6-PanelA and PanelB. The Pareto exponent is around -1 duing the early stage before 1958, then increases from early 1980s to mid 1990s which is also consistant with previous analysis for other samples. This might beacue of the same reason-'Economic Reform' produce more even city size distribution. Then decreases a bit until 2000, after entering the 21st century it remains relatively stable at around -1.8. Just on the whole, each year's absolute value of the Zipf exponent is greater than the top 50 sample, showing that on the whole the top 100 sample is distributed more evernly than top 50 cities over time.

Thirdly, the top 280 cities sample starts from 1984 because there were not so many cities before 1984. The Pareto exponent for these top 280 cities shows closing to V-shape, from 1984 to mid 1990s it constantly increases from -1.17 to -2.32 then decrease a bit until 2009 to -1.91. In other words, the city size distribution for the top 280 large cities is becoming more and more even after 'Economic Reform' to mid 1990s. It is about twice more even than Zipf's prediction, as Zipf's prediction is Pareto exponent equals to -1, therefore, rank 1 city has twice the population of rank 2 city ($(\frac{P_2}{P_1})^{-1} = \frac{r_2}{r_1} = 2$), has three times the population of rank 3 city ($(\frac{P_3}{P_1})^{-1} = \frac{r_3}{r_1} = 3$), etc.. Our Pareto exponent's maximum value is -2.32 (1995), which means that rank 1 city has roughly 1.35 times population of rank 2 city ($(\frac{P_2}{P_1})^{-2.32} = \frac{r_2}{r_1} = 2\ so\ \frac{P_1}{P_2} = \sqrt[2.32]{2} = 1.3482$) and has roughly 1.61 times of population of rank 3 city ($(\frac{P_3}{P_1})^{-2.32} = \frac{r_3}{r_1} = 3\ so\ \frac{P_3}{P_2} = \sqrt[2.32]{3} = 1.6057$). Therefore the top 280 Chinese large cities are distributed more and more evenly within a decade from 1984 to mid 1995s , then a bit less evenly in the next decade from 1995 to 2009. However, during the whole period (1984-2009), the top 280 large cities are distributed much more evenly than Zipf's prediction.

With regard to OLS(2) estimations, the trend of evolution of Zipf's exponents for different subsamples is basically the same with OLS(1) over these 130 years, just the absolute value of Zipf exponent is generally greater than OLS(1) results. While the testing for estimated Pareto exponent= 0 is highly rejected in every year and subsamples (with p-value equals to 0), i.e. the estimated coefficient of Zipf's exponent is significantly different from 0 at 1% significance level, indicating that there indeed is some relationship between rank and size.[30] However, testing for the estimated Pareto exponent= -1 is not significant in all the cases, even for some years the estimated Pareto coefficient is around -1, which implies these estimations are unreliable. The R-square and adjusted R square for each regression is quite high though, around 0.98, but this cannot lead to Zipf's law holding (as Song and Zhang, 2002 conclude the confirmation of Zipf's law from high R-squares), it only shows that the expected linear relationship between rank and size fits the data well.

Therefore, Urzua (2000, 2011) argues the pitfalls of conventional OLS regression of testing for Zipf's law. We calculate the LMZ test statistics as Urzua proposed. Showing in Table3.4 and Table3.5, in all cases, Zipf's law is highly rejected at 10% significance level, as the statistic is much higher than the significance point (4.18 to 6.19). The LMZ value would increase as the number of observations increase, if $(\boldsymbol{z_1}^{\mathbf{2}} + \mathbf{6}\boldsymbol{z_1}\boldsymbol{z_2} + \mathbf{12}\boldsymbol{z_2}^{\mathbf{2}}) > 1$, recall:

$$\mathbf{LMZ} = \mathbf{4n}(z_1^{\mathbf{2}} + \mathbf{6}z_1 z_2 + \mathbf{12}z_2^{\mathbf{2}}) \qquad (9)$$

[30] The testing for non-linearity of E.q.(4) and E.q. (6) is performed for every year and subsamples, as words limit just reported some result for top 50 cities sample in Appendix Table A2.

where

$$\boldsymbol{z_1 \equiv 1 - \frac{1}{n}\sum_{i=1}^{n} ln\frac{x_{(i)}}{x_{(n)}}}$$

and

$$\boldsymbol{z_2 \equiv \frac{1}{2} - \frac{1}{n}\sum_{i=1}^{n} \frac{x_n}{x_i}}$$

So if $(\boldsymbol{z_1}^{\boldsymbol{2}} + \boldsymbol{6z_1z_2} + \boldsymbol{12z_2}^{\boldsymbol{2}}) > 1$, that might be because $\boldsymbol{\frac{1}{n}\sum_{i=1}^{n} ln\frac{x_{(i)}}{x_{(n)}}}$, $\boldsymbol{\frac{1}{n}\sum_{i=1}^{n}\frac{x_n}{x_i}}$ are too small, given that the number of cities n is relatively fixed, so this means the $\boldsymbol{\sum_{i=1}^{n} ln\frac{x_{(i)}}{x_{(n)}}}$ and $\boldsymbol{\sum_{i=1}^{n}\frac{x_n}{x_i}}$ are too small due to the much more even distribution of city sizes. Therefore, LMZ values for Chinese cities are quite large, perhaps because the distribution is much more even than Zipf's law states.

Table 3.3 Pareto exponent for truncated sample

	Top 50 cities			Top 100 cities			Top 280 cities		
Year	OLS(1)	OLS(2)	LMZ	OLS(1)	OLS(2)	LMZ	OLS(1)	OLS(2)	LMZ
1936	-1.199	-1.309	54.825						
1948	-1.133	-1.233	58.610						
1953	-1.114	-1.212	97.313	-1.003	-1.057	208.488			
1958	-1.199	-1.304	78.844	-1.081	-1.140	196.034			
1983	-1.701	-1.857	34.044	-1.451	-1.530	139.686			
1984	-1.825	-1.992	32.516	-1.814	-1.919	85.780	-1.171	-1.200	1133.471
1985	-1.834	-2.001	32.436	-1.941	-2.058	69.314	-1.344	-1.379	738.434
1986	-1.817	-1.982	32.147	-1.882	-1.994	71.695	-1.347	-1.382	716.805
1987	-1.837	-2.004	32.686	-2.033	-2.158	65.460	-1.621	-1.665	490.615
1988	-1.873	-2.045	33.727	-2.096	-2.226	67.880	-1.816	-1.867	315.928
1989	-1.876	-2.050	38.374	-2.111	-2.244	69.551	-1.872	-1.926	315.403
1990	-1.900	-2.076	39.058	-2.150	-2.287	73.482	-1.943	-2.000	309.261
1991	-1.915	-2.093	38.453	-2.164	-2.301	74.271	-1.989	-2.048	300.268
1992	-1.948	-2.131	40.937	-2.203	-2.343	75.716	-2.137	-2.202	242.998
1993	-1.941	-2.129	39.443	-2.223	-2.369	74.079	-2.234	-2.305	202.066
1994	-1.952	-2.140	41.367	-2.273	-2.425	78.619	-2.345	-2.421	194.541

1995	-1.847	-2.019	39.466	-2.170	-2.312	78.869	-2.320	-2.395	194.647
1996	-1.877	-2.052	37.859	-2.190	-2.332	77.314	-2.313	-2.387	192.677
1997	-1.723	-1.881	32.833	-1.788	-1.895	87.511			
1998	-1.587	-1.727	31.138	-1.677	-1.776	88.068			
1999	-1.646	-1.799	32.853	-1.939	-2.065	72.388	-2.153	-2.224	194.748
2000	-1.617	-1.761	31.929	-1.886	-2.004	71.431	-2.093	-2.160	187.444
2001	-1.635	-1.782	31.328	-1.872	-1.989	69.570	-2.065	-2.131	185.766
2002	-1.640	-1.785	31.891	-1.839	-1.951	70.739	-2.006	-2.069	188.733
2003	-1.632	-1.774	31.726	-1.821	-1.931	68.116	-1.978	-2.040	186.703
2004	-1.628	-1.770	32.141	-1.813	-1.922	68.220	-1.971	-2.032	186.242
2005	-1.626	-1.765	32.371	-1.767	-1.869	65.069	-1.913	-1.971	185.050
2006	-1.635	-1.783	31.839	-1.800	-1.910	65.773	-1.939	-1.999	186.117
2007	-1.647	-1.796	32.174	-1.794	-1.903	64.265	-1.922	-1.982	183.897
2008	-1.655	-1.805	32.255	-1.798	-1.908	64.637	-1.918	-1.978	184.351
2009	-1.663	-1.814	32.290	-1.803	-1.912	64.469	-1.908	-1.967	185.365

Notes:

For OLS regressions, Newey-West standard errors controlling for heteroskedasticity reported.

Standard errors, Cook-Weisberg test for heteroskedasticity and Ramsey RESET test for omitted variables see Appendix Table A3.2.

Test for alpha=-1 which is the basic test for Zipf's law and the joint test which is the Alperovich joint test that alpha=1 and Constant= largest city in the basic Zipf regression, also see Appendix Table A3.2.

For OLS(1) and OLS(2) the "*" means significantly different from -1.

Significance points for *LMZ* : 5.98 at 5% level, 4.49 at 10% level.

We choose the top 280 cities in order to capture most of the prefecture-level Chinese cities, up until December 2004 there are 283 prefecture-level cities in total in China.

3.4.3 Panel data testing for Zipf's law

In addition, we also use the panel data to estimate the Pareto exponent, although the previous literature mainly focuses on Pareto exponents for specific years using repeated cross-sectional OLS method. Previous literature mainly focus on the corss sectional data because that Zipf's exponent can be produced for each single year and then makes one is able analyse the evolution of Zipf's law and city size distribution. This also partly because of the lack of available data (they used census data for some specific years). We pool all the available data together and form an unbalanced panel from 1879 to 2009 with the number of cities ranging from 29 to 655.

Then we divide the whole sample into many subsamples to investigate whether Zipf's distribution exists in some specific group of cities. First of all the whole sample is divided into prefecture-level cities and county-level cities. Secondly, in terms of geography, the whole sample is divided into 4 economic regions according to the economic development by Chinese government: East, Midland, West and North-East region cities. Finally, we examine a much longer time series than the previous literature. We extract 4 groups of historical cities from the whole sample according to the political and economic history and available data: (1) Group A. 29 historical cities existing from 1879 to 2009. This sample existed from the end of the 'Qing Dynasty' until now and experienced the war period 1879 to 1949 (WWI, WWII, the Civil War) and the relatively stable period 1948 to 1979 (start of PRC) and the fast development period 1980 to 2009 (after the 'Economic reform'); (2) Group B. 82 cities existing from 1936 to 2009. This sample existed from the end of the first decade of Republic of China- ruled by nationalist party and experienced the same shocks as Group A. Except for WWI. (3) Group C. 125 cities existing from 1958 to 2009. This sample existed from the end of the first decade of the People's Republic of China (the 'new' China) and experienced relatively stable political and economic environment (except for the decade of 'Culture Revolution', 1966-1976) and fast growth period after 'Economic Reform'. (4) Group D. 294 cities existing from 1984 to 2009. This sample existed from the early stage of 'Economic Reform' and experienced a stable political environment and fast growth in every aspect during these 3 decades.

Figure 3.7 and Table 3.4 below describe the Zipf's plots and the OLS regression results for the Pareto exponents for the whole sample and these subsamples respectively. In 4- Panel A, the four columns in each sample's regression refers to the regression equations Eq. (4) to Eq. (7). Analysis will combine the Zipf's plot in Figure 3.7 and estimated Pareto exponent in Table 3.4.

(1) Firstly, the whole sample's Zipf plot shows that Zipf's law does not hold for all the cities during 1879 to 2009. For early years (the relative bottom lines) the Zipf's plot shows a concave shape, then for recent decades as the number of small cities

increasing Zipf's plot is definitely not a straight line with two long lower and upper tails. This is also consistent with the estimated Pareto exponent -0.825 and not significant in testing for equalling to -1. This is also consistent with above results using repeated cross-sectional regression to get the evolution of Pareto exponent in Table 3.2. The value of the estimated Pareto exponent -0.825 indicates that overall Chinese city size distribution is less even than Zipf's prediction.

(2) For the prefectural and county-level cities sample, obviously the prefecture-level cities have the similar Zipf plot with the whole sample and Zipf's distribution cannot approximate the prefecture-level city size distribution, which is consistent with Pareto exponent of -0.872 and significantly different from -1 in Table 3.4. For the county-level cities, Zipf plot shows more concave and flatter shape than prefecture-level cities' and the estimated Pareto exponent of -0.605 also confirms this. County-level cities' distribution is much less even than Zipf's prediction.

(3) Among the 4 economic regions, we find that East cities distribution might follow Zipf's law as estimated Pareto exponent is -1.029 and not significantly different from -1 at 1% level. This is also consistent with above cross-sectional results. The possible explanation is that the East region is historically the most developed area in China especially after the 'Economic Reform' which favours the East most, therefore, overall East region city size distribution is more close to Zipf's law. This is according

to that Zipf's law is mainly proved in developed countries (as shown in the literature section), and more developed regions seem to have more balanced development of all cities (the disparity between large and small cities is not too large or too small). [31]

For cities in other regions, Midland cities are distributed the least evenly; West cities and North-East cities are all distributed a bit less even than Zipf's distribution.

(4) Among the 4 historical cities group, Zipf's law is found in Group B- 82 cities, 1936 to 2009 with Pareto exponent -0.977 (corrected OLS result) which is not significantly different from -1. The Zipf's plot is also close to a straight line with slope -1 as Zipf's law states. This shows that these historical cities are distributed as even as Zipf's law prediction that ranking 1 city has roughly twice the population of the 2nd ranked city and has roughly three times the population of the 3rd ranked city etc. Other groups show an increasing Pareto exponent over time from Group A to Group D. The oldest cities- 29 (1879 to 2009) - are distributed less evenly than Zipf's distribution; Group B- 82 cities (1936-2009) are distributed just as Zipf's prediction; then Group C- 125 cities (1958-2009) are distributed more evenly than Zipf's law; the younger city group from 1984 to 2009 obviously includes a lot small cities according to Zipf's plot and shows more evenness than Zipf's law with a Pareto exponent -1.3. This is consistent with the previous cross-sectional results of evolution of Pareto

[31] There are still no exact studies about the level of economic development or the level of income and the validity of Zipf's law.

exponent that cities are distributed more and more evenly over time, Zipf's law shows in the middle of this process.

In addition, model 2 and model 3 show whether the quadratic term and cubic term should be included in the model. Generally, the estimated coefficient for the quadratic term in every subsample is significantly different from 0 at 1% significance level, which indicates the existence of nonlinearity of the relationship between city rank and size which is also showed in Zipf plots (not straight line, but a bit concave). Model 4 uses the corrected OLS model (Gabaix and Ibragimov, 2011) as showed in Eq. (7), except for the absolute value of Pareto exponent which is a little bit greater than model 1 (conventional OLS), the significance and trend are the same.

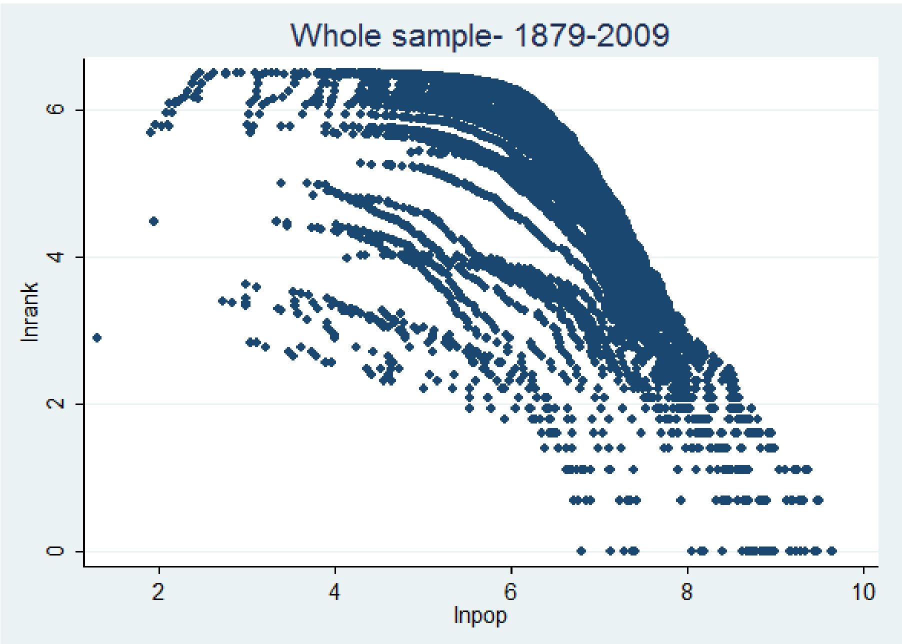


Figure 3.7_Panel A: Zipf's plot for panel data- whole sample.

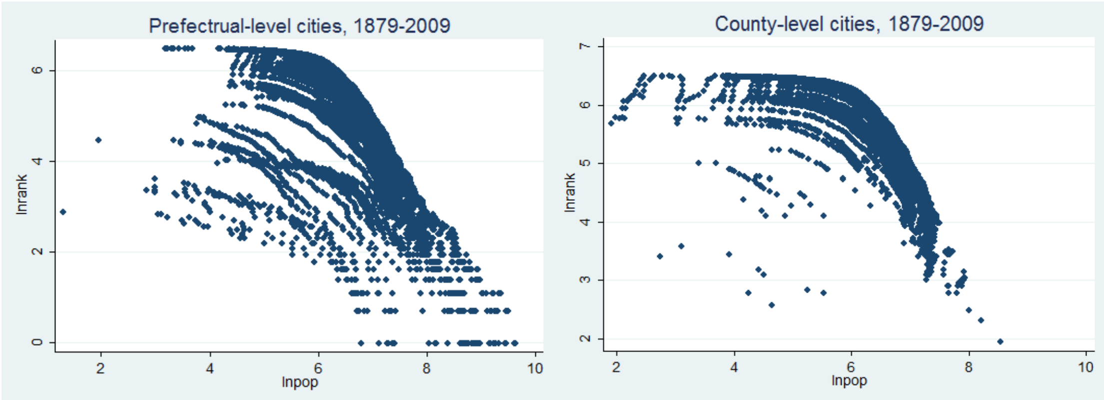


Figure 3.7_Panel B: Zipf's plot for panel data- prefecture-level and county-level cities.

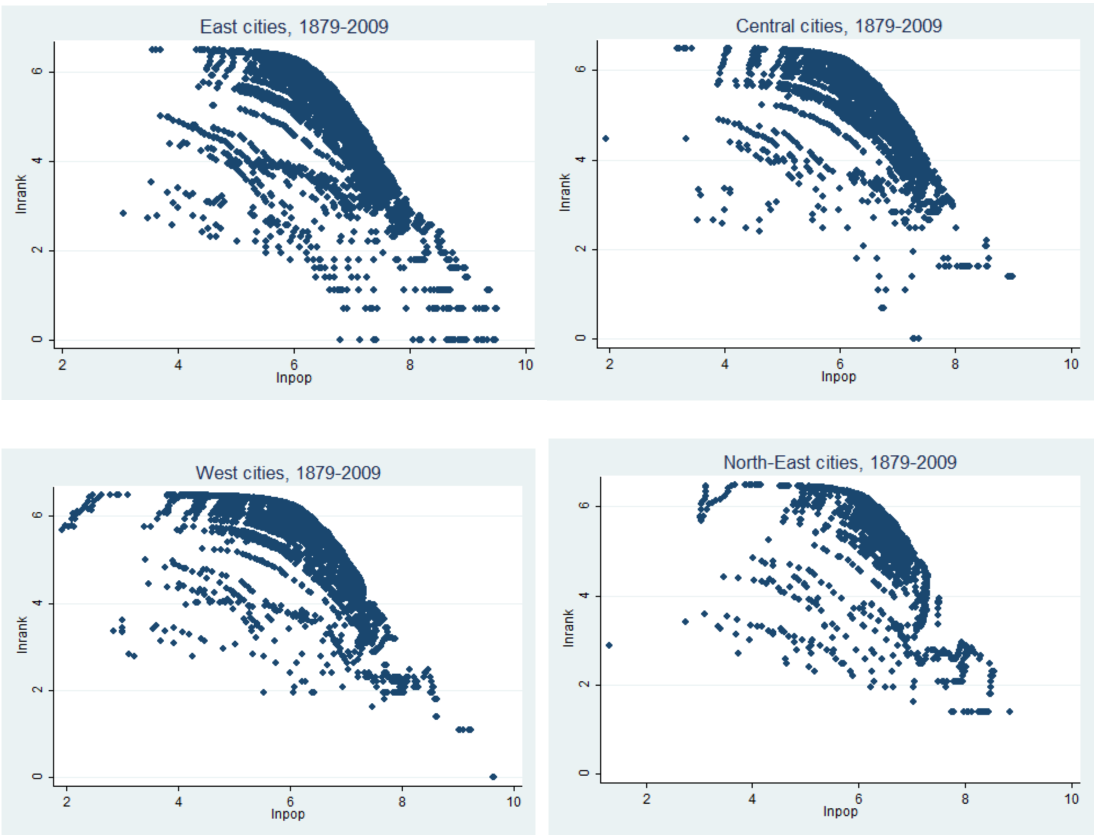


Figure 3.7_Panel C: Zipf's plot for panel data- regional sample.

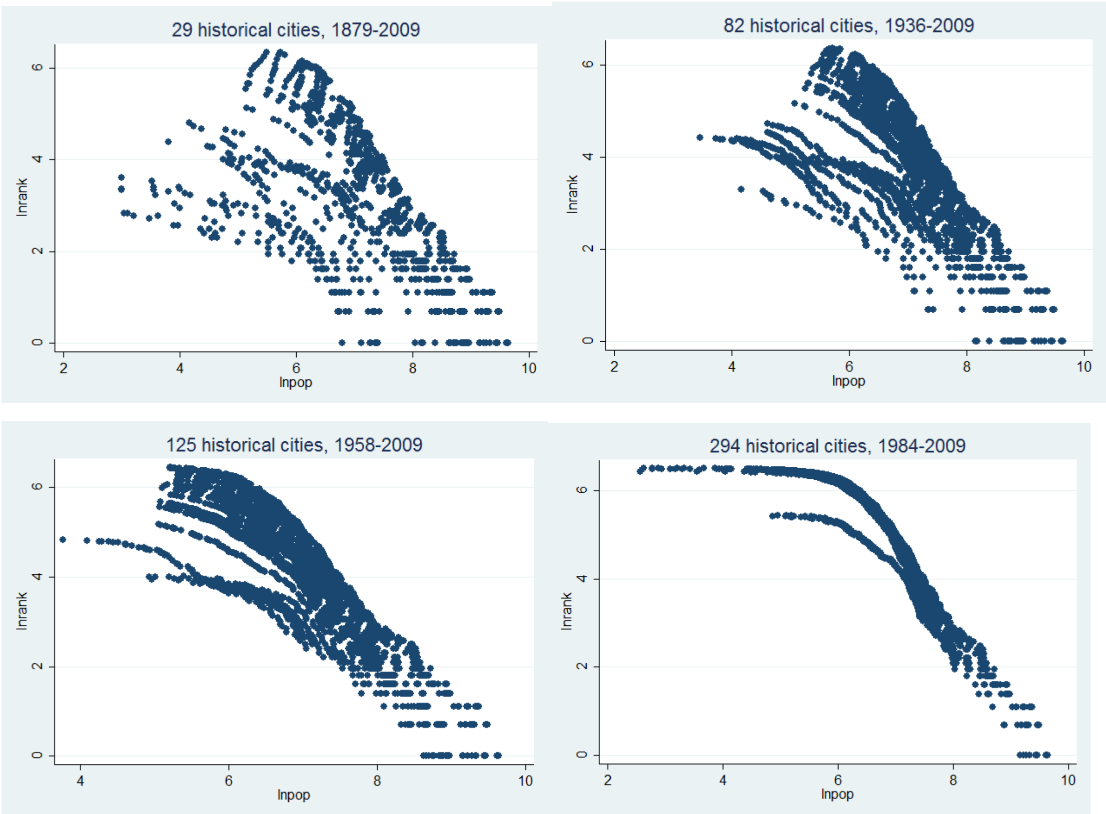


Figure 3.7_Panel D: Zipf's plot for panel data- historical cities sample.

Table 3.4_ Panel A: OLS for Pareto exponent on unbalanced panel data.

	Whole sample			
	(model 1)	(model 2)	(model 3)	(model 4)
	lnrank	lnrank	lnrank	lnrank2
lnpop	-0.825***	3.859***	1.524***	-0.842***
	(0.0159)	(0.125)	(0.485)	(0.0166)
$(\text{lnpop})^2$		-0.381***	0.0216	
		(0.00963)	(0.0798)	
$(\text{lnpop})^3$			-0.0223***	
			(0.00433)	
Constant	10.42***	-3.707***	0.633	10.52***
	(0.103)	(0.401)	(0.976)	(0.107)
Observations	14,999	14,999	14,999	14,999
R-squared	0.367	0.584	0.587	0.367
Test $\alpha = -1$	0.0000	0.0000	0.0000	0.0000

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 In Stata 12, using *reg y x, vce(robust)* command. *vce(robust)* uses the robust or sandwich estimator of variance, obtaining robust variance estimates.

Table 3.4 _Panel B: OLS for Pareto exponent on unbalanced panel data- prefectural and county-level cities sample.

	Prefecture-level cities				County-level cities			
	(model 1)	(model 2)	(model 3	(model 4)	(model 1)	(model 2)	(model 3)	(model 4)
VARIABLES	lnrank	lnrank	lnrank	lnrank2	lnrank	lnrank	lnrank	lnrank2
lnpop	-0.872***	4.817***	7.312***	-0.899***	-0.605***	3.057***	-5.204***	-0.607***
	(0.0227)	(0.179)	(1.862)	(0.0237)	(0.0155)	(0.143)	(0.347)	(0.0156)
lnpop2		-0.442***	-0.848***			-0.321***	1.301***	
		(0.0135)	(0.287)			(0.0118)	(0.0682)	
lnpop3			0.0214				-0.102***	
			(0.0144)				(0.00433)	
Constant	10.42***	-7.538***	-12.49***	10.58***	9.381***	-0.806*	12.47***	9.390***
	(0.152)	(0.587)	(3.946)	(0.158)	(0.0978)	(0.427)	(0.572)	(0.0982)
Observations	7,810	7,810	7,810	7,810	7,189	7,189	7,189	7,189
R-squared	0.343	0.565	0.568	0.345	0.455	0.708	0.775	0.454
Test $\alpha = -1$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 3.4- Panel C: OLS for Pareto exponent on unbalanced panel data- 4 economic regions sample.

	East cities				North-East cities			
	(model 1)	(model 2)	(model 3	(model 4)	(model 1)	(model 2)	(model 3)	(model 4)
VARIABLES	lnrank	lnrank	lnrank	lnrank2	lnrank	lnrank	lnrank	lnrank2
lnpop	-1.029***	5.433***	15.27***	-1.063***	-0.686***	4.237***	-1.021	-0.695***
	(0.0299)	(0.194)	(1.410)	(0.0316)	(0.0478)	(0.273)	(0.857)	(0.0485)
lnpop2		-0.491***	-2.011***			-0.412***	0.533***	
		(0.0147)	(0.208)			(0.0212)	(0.150)	
lnpop3			0.0766***				-0.0544***	
			(0.0101)				(0.00854)	
Constant	11.74***	-9.257***	-30.03***	11.95***	9.508***	-4.864***	4.390***	9.557***
	(0.200)	(0.642)	(3.152)	(0.211)	(0.308)	(0.876)	(1.603)	(0.312)
Observations	5,843	5,843	5,843	5,843	2,033	2,033	2,033	2,033
R-squared	0.404	0.609	0.622	0.405	0.249	0.520	0.537	0.250
Test $\alpha = -1$	0.3319	0.0000	0.0000	0.0453	0.0000	0.0000	0.9802	0.0000

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

In Stata 12, using *reg y x, vce(robust)* command. *vce(robust)* uses the robust or sandwich estimator of variance, obtaining robust variance estimates.

Table 3.4- Panel C (continue): OLS for Pareto exponent on unbalanced panel data- 4 economic regions sample.

	Midland cities				West cities			
	(model 1)	(model 2)	(model 3	(model 4)	(model 1)	(model 2)	(model 3)	(model 4)
VARIABLES	lnrank	lnrank	lnrank	lnrank2	lnrank	lnrank	lnrank	lnrank2
lnpop	-0.750***	5.244***	2.984	-0.758***	-0.722***	2.939***	0.435	-0.731***
	(0.0334)	(0.365)	(2.161)	(0.0339)	(0.0251)	(0.159)	(0.579)	(0.0258)
lnpop2		-0.490***	-0.105			-0.313***	0.149	
		(0.0289)	(0.354)			(0.0128)	(0.109)	
lnpop3			-0.0214				-0.0272***	
			(0.0191)				(0.00672)	
Constant	10.01***	-8.024***	-3.720	10.06***	9.781***	-0.629	3.651***	9.832***
	(0.215)	(1.148)	(4.344)	(0.218)	(0.155)	(0.487)	(1.009)	(0.159)
Observations	5,843	5,843	5,843	5,843	2,033	2,033	2,033	2,033
R-squared	0.404	0.609	0.622	0.405	0.249	0.520	0.537	0.250
Test $\alpha = -1$	0.3319	0.0000	0.0000	0.0453	0.0000	0.0000	0.9802	0.0000

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

In Stata 12, using *reg y x, vce(robust)* command. *vce(robust)* uses the robust or sandwich estimator of variance, obtaining robust variance estimates.

Table 3.4_ Panel D: OLS for Pareto exponent on balanced panel data- 4 historical cities sample.

	Group-A 29 cities, 1879-2009				Group-B 82 cities, 1936-2009			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	lnrank	lnrank	lnrank	lnrank2	lnrank	lnrank	lnrank	lnrank2
lnpop	-0.752***	3.556***	4.954***	-0.809***	-0.931***	4.717***	14.29***	-0.977***
	(0.0385)	(0.183)	(1.128)	(0.0413)	(0.0271)	(0.154)	(1.681)	(0.0291)
lnpop2		-0.316***	-0.537***			-0.413***	-1.854***	
		(0.0134)	(0.173)			(0.0114)	(0.244)	
lnpop3			0.0112				0.0706***	
			(0.00855)				(0.0116)	
Constant	8.498***	-5.711***	-8.531***	8.825***	10.34***	-8.596***	-29.29***	10.62***
	(0.289)	(0.610)	(2.373)	(0.308)	(0.195)	(0.515)	(3.792)	(0.207)
Observations	1,156	1,156	1,156	1,156	2,828	2,828	2,828	2,828
R-squared	0.321	0.489	0.490	0.328	0.423	0.607	0.619	0.429
Test $\alpha = -1$	0.0000	0.0000	0.0000	0.0000	0.0111	0.0000	0.0000	0.4347

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

In Stata 12, using *reg y x, vce(robust)* command. *vce(robust)* uses the robust or sandwich estimator of variance, obtaining robust variance estimates.

Table 4_ Panel D (continue): OLS for Pareto exponent on balanced panel data- 4 historical cities sample

	Group-C 125 cities, 1958-2009				Group-D 294 cities, 1984-2009			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	lnrank	lnrank	lnrank	lnrank2	lnrank	lnrank	lnrank	lnrank2
lnpop	-1.325***	3.403***	22.87***	-1.372***	-1.300***	2.612***	5.809***	-1.326***
	(0.0209)	(0.193)	(1.921)	(0.0226)	(0.0202)	(0.0829)	(0.740)	(0.0214)
lnpop2		-0.339***	-3.169***			-0.295***	-0.804***	
		(0.0138)	(0.270)			(0.00656)	(0.110)	
lnpop3			0.135***				0.0263***	
			(0.0125)				(0.00537)	
Constant	13.42***	-2.842***	-46.76***	13.72***	13.75***	1.021***	-5.475***	13.91***
	(0.148)	(0.666)	(4.502)	(0.159)	(0.136)	(0.260)	(1.637)	(0.144)
Observations	3,722	3,722	3,722	3,722	4,358	4,358	4,358	4,358
R-squared	0.655	0.726	0.749	0.658	0.817	0.924	0.928	0.813
Test $\alpha = -1$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

In Stata 12, using *reg y x, vce(robust)* command. *vce(robust)* uses the robust or sandwich estimator of variance, obtaining robust variance estimates.

Furthurmore, we could also see the evidence of rejecting Zipf's law at the upper tail, directly from Table 3.5 panel A to C which presents the 20 largest cities' population and the value that rank 1 city's population over each city's population, from 1990 to 2009. As Zipf's law states that city size is perfectly inversely correlated to its rank, with exponent equal to -1:

$$\mathbf{r} = \mathbf{c}\boldsymbol{p}_{(r)}^{-\mathbf{1}}$$

then

$$\boldsymbol{p}_{(r)} = (c/r)^1$$

thus, $\frac{\boldsymbol{p}_{(1)}}{\boldsymbol{p}_{(2)}} = \frac{c/1}{c/2} = \mathbf{2}$, and therefore

$$\frac{\boldsymbol{p}_{(1)}}{\boldsymbol{p}_{(r)}} = \frac{\boldsymbol{c/1}}{\boldsymbol{c/r}} = \frac{\boldsymbol{r}}{\mathbf{1}}$$

If there indeed exists a power law relationship with Pareto exponent -α , but Zipf's law does not hold, then

$$\mathbf{r} = \mathbf{c}\boldsymbol{p}_{(r)}^{-\boldsymbol{\alpha}}$$

$$\boldsymbol{p}_{(\boldsymbol{r})} = (c/r)^{\frac{1}{\alpha}}$$

$$\frac{\boldsymbol{p}_{(\mathbf{1})}}{\boldsymbol{p}_{(\boldsymbol{r})}} = (\frac{\boldsymbol{c}/\mathbf{1}}{\boldsymbol{c}/\boldsymbol{r}})^{\frac{1}{\alpha}} = \boldsymbol{r}^{\frac{1}{\alpha}}$$

which means that rank 1 city will have $\mathbf{2}^{\frac{1}{\alpha}}$ times the population of rank 2 city; have $\mathbf{3}^{\frac{1}{\alpha}}$ times the population of rank 3 city, etc. So if Zipf's law is true for China, the population of the first largest (rank1) city should be tiwce as large as the second largest city and threetimes bigger than the third largest city, so forth, be the r multiple bigger than the rank r city. Obviously there is no sign of Zipf's law in Chinese city size in Table 3.7. But there indeed exists a Power law in Chinese city hierarchy, that city size distribution follow a power law with exponent greater than 1 (absolute value), which also indicates that Chinese city size distribution for large cities is more equal (even) than Zipf's law prediction. This is contrary to previous literature in developed countries that Zipf's law holds at least for large cities (Rosen and Resnick, 1980; Parr, 1985; Soo, 2005; Rose, 2005; Krugman, 1996; Eatonand Eckstein, 1997; Gabaix, 1999a; Dobkins and Ioannides, 2000; Davis and Weinstein, 2002; Brosker *et al.*, 2008; Berry and Okulica-Kozaryn, 2012) mainly because China's unique history and economy development process, large cities emerge quite early thus has long been the population agglomeration area and are developed primarily with favorable policies, the disparity between large cities is reasonablely to be small (distribute more equally).

The results are also contrary to early studies on Chineses cities (Song and Zhang, 2002; Gan, Li and Song, 2006), which is mainly because early studies on China focus on only two or three specific years. As previous Pareto evolution results show that, Chinese city size distribution is becoming more and more even from 1879 to 2009, this process includes Zipf's law for some years (1988-1992).

Table 3.5 _ Panel A Twenty Largest Cities in China, 2009

2009			
rank	city	population p(r) (r=1,..,20)	p(1)/p(r)
1	Chongqing	15427.7	1.0000
2	Shanghai	13316.8	1.1585
3	Beijing	11746.3	1.3134
4	Tianjin	8029	1.9215
5	Guangzhou	6546.8	2.3565
6	Xi'an	5615.8	2.7472
7	Nanjing	5459.7	2.8257
8	Chengdu	5208.6	2.9620
9	Wuhan	5149.7	2.9958
10	Shenyang	5122.3	3.0119
11	Shantou	5034.3	3.0645
12	Harbin	4747	3.2500
13	Hangzhou	4294.4	3.5925
14	Foshan	3676.3	4.1965
15	Changchun	3623.2	4.2580
16	Jinan	3482.4	4.4302
17	Tangshan	3070	5.0253
18	Dalian	3020.1	5.1083
19	Taiyuan	2851.6	5.4102
20	Zhengzhou	2850.1	5.4130

Panel B Twenty Largest Cities in China, 2000

2000			
rank	city	population x(r) (r=1,..,20)	p(1)/p(r)
1	Shanghai	11368.2	1
2	Beijing	9741.4	1.1669986
3	Chongqing	8964.9	1.2680788
4	Wuhan	7491.9	1.5173988
5	Tianjin	6820.5	1.6667693
6	Guangzhou	5666.8	2.0061057
7	Huaian	5102.7	2.2278794
8	Shenyang	4850.4	2.3437655
9	Xi'an	3934.7	2.8892165
10	Chengdu	3358.6	3.3848032
11	Harbin	3037.2	3.742987
12	Jinmen	2988	3.8046185
13	Changchun	2928.3	3.8821842
14	Nanjing	2895.2	3.9265681
15	Zibo	2685	4.2339665
16	Dalian	2677.8	4.2453507
17	Jinan	2644.6	4.2986463
18	Chaoyang	2494	4.5582197
19	Qingdao	2346	4.8457801
20	Taiyuan	2332	4.8748714

Panel C Twenty Largest Cities in China, 1900

1990			
rank	city	population x(r) (r=1,..,20)	p(1)/p(r)
1	Shanghai	7834.8	1
2	Beijing	6995.1	1.1200412
3	Tianjin	5771	1.3576157
4	Shenyang	4538.7	1.7262212
5	Wuhan	3750.5	2.0890015
6	Guangzhou	3579.4	2.1888585
7	Chongqing	2984.4	2.6252513
8	Harbin	2827.1	2.7713204
9	Chengdu	2808.1	2.7900716
10	Xi'an	2756.7	2.8420938
11	Nanjing	2497.5	3.1370571
12	Zibo	2457.5	3.188118
13	Dalian	2396.4	3.2694041
14	Jinan	2322.7	3.3731433
15	Changchun	2110	3.7131754
16	Qingdao	2057.8	3.8073671
17	Taiyuan	1964.3	3.9885964
18	Liupanshui	1826.9	4.2885763
19	Zaozhuang	1775.4	4.4129774
20	Zhengzhou	1705.6	4.5935741

Source: *China Urban Statistical Yearbook*, 2010. Population unit: 1000 person.

To conlude, firstly we do graphical analysis to get the Zipf plot (we plot lnrank on logsize), and Zipf's law may hold in the middle ranked cities where the plot seems to show a stright line. And then we perform the conventional OLS estimation of logarithm of rank on the logarithm of size, it is repeated cross sectional regression because the rank is valid only in each individual year. We test for the whole sample and prefecture-level cities, county-level cities, Top 50 cities, Top 100 cities, Top 280 cities subsamples, the results show that generally Zipf's law does not hold for most of the years. We only find Zipf's law emerged for a few years during the process of urban growth at the end of 1980s and the early 1990s (the Zipf exponent is around -1 from 1989 to 1992, but testing for Zipf exponent equals to -1 is rejected in all cases). Results also show that prefecture-level cities are distributed more evenly than county-level cities which may be due to the policy priorities; during the first two decades of 'Economic Reform' (1980s and 1990s) the top 50 cities and top100 cities (large cities) are particulaly distributed more evenly than Zipf's law predicts.

Since the OLS estimation is biased, we emply the new method of LM test, which shows that Zipf's law does not hold in Chinese cities from 1879 to 2009, even in upper truncated samples (top 50 subsample, the large cities). This is contrary to the previous literature which suggests that Zipf's law holds at least in the upper truncatd tail (see Rosen and Resnick, 1980; Ioannides and Overman, 2003; Soo, 2004 for US and world wide case; Song and Zhang, 2002; Schaffar, 2010 and Schaffar and Dimou, 2012 for Chinese cases).

In additon, we also perform panel data estimation for Zipf exponent. Four models are constructed, including conventional OLS, models with quadratic term and cubic term, models using corrected OLS. Then test for the subsample of prefecture-level cities, county-level cities, East cities, North-East cities, Midland cities and West cities, we found that the distribution of East cities could support Zipf's law as Zipf exponent equals to -1.029, but the test for Zipf exponent equals to -1 is still insignificant.

3.4.4 Urban primacy measure versus Zipf's exponent

While urban economists focus mainly on the Pareto to measure the size distribution of cities, development economists tend to focus on measure of primacy to investigate the urban system.

Rosen and Resnick (1980) examine and compare the Pareto and primacy measure of the size distribution of cities for a sample of 44 countries. There is the Williamson (1965) hypothesis, as adapted to an urban context (Hansen, 1990), which states that a high degree of spatial or urban concentration in the early stages of economic development is helpful, through spatially concentrating industrialisation. As the development proceeds, eventually deconcentration of primacy occurs, possibly because that firstly the economy can afford to spread economic infrastructure and knowledge resources, secondly, the cities of initial high concentration become high cost, congested. There is a number of studies find the pattern of primacy first increasing and then decreasing as income rises over time (EI-Shakhs, 1972; Alonso, 1980; Junius, 1999; Davis and Henderson, 2003). In addition, Henderson (2003) argues that the form of urbanisation or the degree of urban concentration strongly affects productivity growth. For any country size and level of development, there is a best degree of urban concentration, which balances the gains from enhanced concentration (such as local knowledge accumulation) against the losses caused by over concentration (such as congestion). The best urban primacy declines with larger country size and higher level of development. This also supports our analysis of policy impact of 'Economic Reform' on city size distribution previously. 'Economic Reform' promotes the level of economics and development. Hence, it produces more even city size distribution.

Primacy can be defined in various ways, [32] generally, the principle is that primacy is to measure the extent to which the largest city or cities dominate the country's urban hierarchy. According to Henderson (2003), any city's share of national urban population should decline as national urban population grows and more cities form. We measure primacy by urban population rather than total population.

---

[32] Such as the ratio of the largest city to the first two, three, four, or more cites.

One can argue that changes in urban primacy would then translate into changes in Zipf's law. We would expect Zipf's exponent has a negative correlation with measure of primacy, because higher Zipf's exponent indicates more evenly distributed city size (population) where one will observe a lower primacy, as Rosen and Resnick (1980) confirm this negative relationship using a sample of 44 countries and also confirm that Zipf's exponent is a better reflection of the overall city-size distribution than is primacy measure. We simply display a bit of this kind of measurement of urban development to illustrate what happens to urban primacy of the large city (or cities) over time, and also to compare with Zipf's exponent.

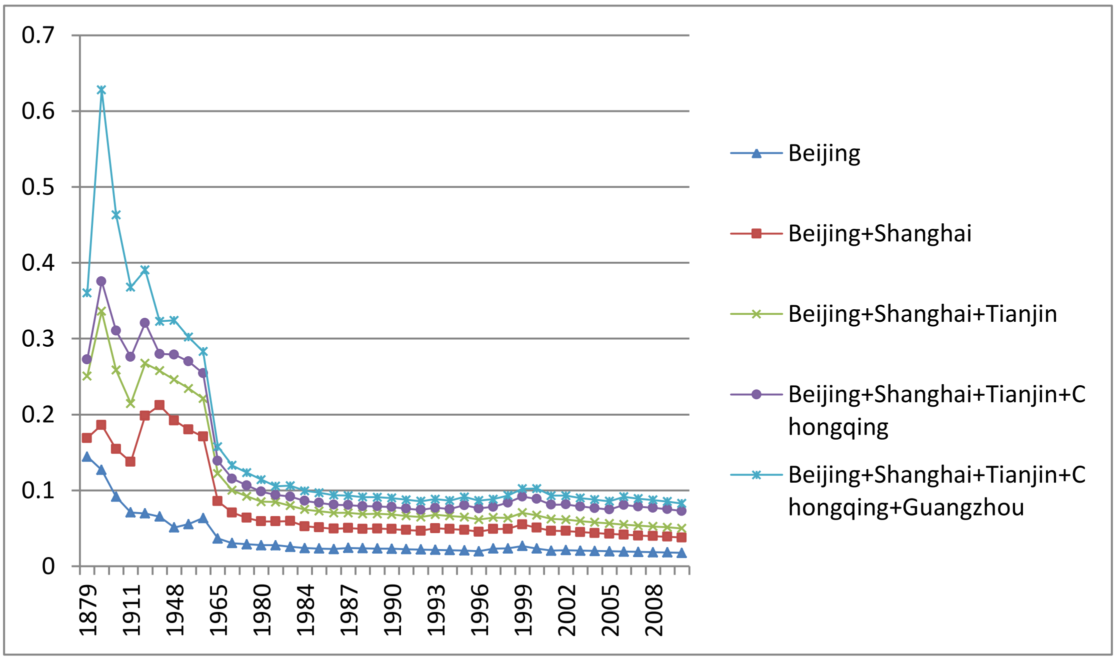


Figure 3.8 urban primacy over time (the proportion of population of largest cities on total urban population for each year)

As shown in Figure 3.8, generally speaking, as time passes urban primacy decreases. This is consistent with Zipf's exponent which generally increases over time, indicating that city size distribution become more evenly, thus the proportion of large cities of overall urban population will decrease. Specially, as early as 1900s, the end of Qing Dynasty, large cities dominates the urban system. The five largest cities Beijing, Shanghai, Tianjin, Chongqing, and Guangzhou accounts for over 60% of total urban population, and we can also see that Guangzhou was relatively largest as if one expect for Guangzhou and the four largest cities accounts for less than 40%. This is consistent with the low Zipf's exponent around 0.6 to 0.7 during that period (Table 3.2).

And then urban primacy experienced sharply decrease around 1920s, this might because of the establishment of Republic of China in 1912 which may lead to the increase of number of cities. There is another sharp decrease of urban primacy around 1960s. The possible explanations might be the establishment of People's Republic of China in 1949, which cause a large number of cities emerge and the growth of mediate and small cities. These two sharp decrease consistent with the increase of Zip's exponent in 1920s and 1930s (Table 3.2).

After 1960s, the urban primacy gradually decreases a bit and then relatively stable after mid-1980s. This might because after mid-1980s 'Economic Reform 1979' promotes the whole economy substantially and all cities grow identically fast, thus urban primacy stable and also consistent with Zipf's exponent (around 1, from 0.97 to 1.18).

3.5 Conclusion

In this paper, we investigate the evolution of Chinese city size distribution by testing for the well-known regularity- Zipf's law- from 1879 to 2009. Zipf's law emerged in the middle of the process of urban growth during this period. Specifically, Chinese city size distribution evolves from less even than Zipf's prediction (before 'Economic Reform'[33], 1983) to very close to Zipf's prediction (end of 1980s and early 1990s) then grows to more even than Zipf's law prediction, then relatively stable with Pareto exponent around -1.15. This process is consistent with the literature trying to explain Zipf's law that Zipf's law will emerge under homogeneous urban growth, but city size needs time to converge to Zipf's law (Gabaix, 1999), but no previous literature has covered the long run evolution of Chinese city size distribution. In addition, we divide the whole sample into many subsamples to investigate whether Zipf's distribution exists in some specific group of cities: firstly prefecture-level cities and county-level cities; secondly, 4 economic regions: East, Midland, West and North-East cities; thirdly, 4 groups of historical cities: 29 cities (1879-2009), 82 cities (1936-2009), 125

[33] In this paper, 'Economic Reform' refers to the year that it becomes effective, like early stage of 1980s: 1983, 1984, 1985 etc.

cities (1958-2009) and 294 cities (1984-2009). Zipf's law is found in prefecture-level cities, East region cities and 82 cities from 1936 to 2009.

Specifically, firstly we find prefecture-level cities have a similar distribution evolution as the whole sample, but after the 'Economic Reform' they begin to grow much more equally than Zipf's prediction (disparity between large and small cities is smaller). This might because that 'Economic Reform' promote the economic development and enhance the overall income levels, which decrease the differences between city sizes. After 2005 the distribution is relatively stable with a Pareto exponent of -1.5; while county-level cities have long been distributed less evenly than Zipf's prediction (disparity between large and small cities is quite large), but Pareto exponent shows a trend to Zipf's law. This might be because in the beginning county-level cities are always not the policy priority (but some particular county-level cities that have a long history have been picked out for a nice development[34]).

Subsequently, the distribution for county-level cities grows rapidly toward an even distribution; especially after the 'Economic Reform' county-level cities receive more care. After 2000, the distribution of county-level cities tends to be stable with Pareto exponent around -0.82 which close to but still not as even as Zipf's law. Secondly, Zipf's law is found in East region cities with a panel OLS regression Pareto exponent

[34] For instance, as a county-level city Puning City (in Guangdong Province) ranked 35th in 2009 by population, which is higher than most of the prefecture-level cities.

-1.029 and testing for $\alpha = -1$ cannot be rejected at 1% significant level. Because East region has long been the most developed region in China and has the policy priority for development especially in early stage of 'Economic Reform', which means that the East region is growing under a homogeneous urban growth process and produces Zipf's distribution on the whole. This is supportive of Gabaix's (1999) theory that Zipf's law comes from stochastic homogeneous urban growth. Thirdly, Zipf's law is found in the group of 82 older cities from 1936 to 2009.

While previous studies of developed countries often find that Zipf's law holds at least for large cities, here Zipf's law is often rejected in upper tail, i.e. for large cities in China. Large Chinese cities tend to be distributed much more equally than Zipf's law predicts with the absolute value of the Pareto exponent constantly significantly greater than 1 (except for top ranking 100 cities for the years 1953 and 1958). However, the OLS estimation and Graphs show that the medium-sized cities might follow Zipf's law.

CHAPTER 4
CHINESE CITY SIZE DISTRIBUTION: TESTING FOR GIBRAT'S LAW

4.1 INTRODUCTION

Gibrat's law states that the growth rate of urban population is independent of its initial size, which also indicates that all cities will grow with the same expected rate and the same variance, regardless of their various sizes. This is another stylized fact that addresses the local population growth and can explore the city size distribution dynamically. Basically, Zipf's law in previous chapter and Gibrat's law are the two sides of the same coin, Gibrat's law is about the city growth process and Zipf's law is exploring the resulting population distribution.

Gibrat's law is also known as the 'rule of proportionate growth' and relates to city growth and city size. Originally, it refers to the increment of city growth being proportional to its initial size like $x_t - x_{t-1} = \varepsilon_t x_{t-1}$, Steindl (1965), where x_t is the size of the city at time t and ε_t is the proportional rate of growth between period (t-1) and t. Then after some derivation, finally we have that the distribution of $\log x_t$ is approximated by a normal distribution of $\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_t$, with mean μt and variance σ^2 (assuming that the proportion rate ε_t to be independent variates with mean μ and variance σ^2). Therefore, the urban growth pattern is described as a homogeneous urban growth process with a common mean and variance. More simply, Gibrat's law also indicates that the subsequent growth rate of a city is independent of its initial size as $\frac{x_t - x_{t-1}}{x_{t-1}} = \varepsilon_t$, and the process behind city growth is stochastically growth proportionate to its size. In other words, the underlying stochastic process is the same for all cities, and large cities should not grow faster than cities of other sizes.

In addition, the results of testing for Gibrat's law will further support our finding of Zipf's law as Gibrat's law will deliver Zipf's law at the steady state of Gibrat's mode of growth (Gabaix 1999). We find that Gibrat's law gradually emerged although it was not fully attained, which could explain the results for Zipf's distribution. However, unlike Gabaix's (1999) argument which states that Zipf's law will emerge at the steady state, in our study the Chinese city size distribution firstly shows less evenness than Zipf's law predicts, and then it is consistent with Zipf's law predicts for a few years, finally it becomes more even (equal) than Zipf's prediction, i.e. the disparity between large and small cities is smaller and smaller during 1879 to 2009 and finally grows beyond Zipf's law prediction.

Data we used is the same as in the study of Zipf's law, it is Chinese city level population data from 1879 to 2009 and the number of cities for each year varies from around 30 to over around 600. We also grouped the data into several subsamples. In prefecture-level cities, Gibrat's law also holds which offers the explanation of the emergence of Zipf's law. For county-level cities, Gibrat's law does not hold. The correlation coefficient between city growth rate and initial size is consistently negative. This indicates that in the county-level cities' sample, large cities grow a bit slower than smaller ones, which is consistent with the convergence trend of county-level city growth (absolute Pareto exponent increases over time, showing city size distribution more and more equal). For the different regions, Gibrat's law is found in the East region, which indicates that East region cities grow stochastically with a homogeneous growth process, and then Zipf's law emerged for the city size distribution. Other regions do not support Gibrat's law. The group of 82 historical cities (1936-2009) tends to be the closest one toward Gibrat's law. While other historical cities show the evolution of growing toward to Gibrat's mode of urban growth.

4.2 LITERATURE REVIEW

4.2.1 Theoretical Review

Gibrat's law is involved in the study of city size distribution, since Gabaix (1999) explains Zipf's law using Gibrat's law (another well-known empirical regularity). The original Gibrat's law is proposed by Gibrat (1931) who presents the first formal model of the dynamics of firm size and announced a 'new law': the Law of Proportional Effect. He assumes that the underlying process of widespread skewed distributions in various areas is a simple Gaussian process: a large number of small additive influences, operating independently of each other, would generate a normally distributed variablez. Therefore an observed skewed distribution of some variable x could be modeled by assuming that some underlying function of x was normally distributed. Gibrat (1931) postulates the 'simplest' such process by assuming that the logarithm of x was normally distributed. And the 'Proportional growth' means the expected value of an increment to a firm size in each period is proportional to the current size of the firm. Following Steindl (1965) the argument can be presented as

$$x_t - x_{t-1} = \varepsilon_t x_{t-1}$$

where x_t is the size of the firm at time t and ε_t is the proportional rate of growth between period (t-1) and t. Therefore,

$$x_t = (1 + \varepsilon_t)x_{t-1} = x_0(1 + \varepsilon_1)(1 + \varepsilon_2) \dots (1 + \varepsilon_t)$$

If we choose a 'short' time period, then we can regard ε_t as being 'small', so $\log(1 + \varepsilon_t) \approx \varepsilon_t$. Then taking logarithm of above equation we obtain:

$$\log x_t \approx \log x_0 + \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t$$

By assuming the proportion rate ε_t to be independent variables, with mean μ and variance σ^2, and as $t \to \infty$, $\log x_0$ would be small comparing to $\log x_t$. Finally, we have that the distribution of $\log x_t$ is approximated by a normal distribution of $\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t$, with mean μt and variance σ^2. *"The law of proportionate effect will therefore imply that the logarithms of the variable will be distributed following* *the normal distribution"—Robert Gibrat (1931).*

Gabaix (1999) applies Gibrat's law in cities which indicates that city size distribution tends to be log-normal and his explanation was that the growth process can be multiplicative and independent of size. The growth rate is

$$\frac{x_t - x_{t-1}}{x_{t-1}} = \varepsilon_t$$

which is obviously independent of initial size. In other words, even though growth rates between different cities can vary substantially, there is no systematic pattern with respect to size, i.e., the underlying stochastic process is the same for all cities. This proposition also attracts numerous empirical studies testing for its validity, and normally together with testing for Zipf's law. This leads to a debate in various fields including urban economics, urban geography and statistical physics about which laws could approximate the city size distribution better Pareto distribution or log-normal distribution. And what are the appropriate urban units to which the size distribution model should be fitted?

4.2.2 Empirical Review

The independence of the city growth rate and its initial size implies that the approximate distribution of city population will be log-normally distributed across all cities (Gabaix, 1999; Eeckhout, 2004). In terms of the empirical evidence, for different countries and various time sapns, using different methods, whether Gibat's law holds is still controversial.

On the one hand, Gibrat's law has been observed within many countries at different time periods. Firstly for U.S. cities, as early as 1956, Madden finds that the distribution model of different-sized cities was quite stable, though cities expand quickly during the period 1790 to 1950, which implies that Gibrat's law holds. Glaeser *et al.* (1995) support Gibrat's law and find that there is no difference in the mean growth rate between large and small cities from 1950 to 1990. Gabaix (1999a) considers a fixed number of cities growing stochastically. He uses the U.S. 135 large metro areas for 1991 and finds that Gibrat's law holds. There exists homogeneity of growth process, i.e. cities follow similar processes with a common mean and variance. Ioannides and Overman (2003) use U.S. census data for metro areas from 1900 to 1990 and find that despite variations in growth rates of cities Gibrat's law does hold. They non-parametrically estimate a stochastic kernel – a three-dimensional representation of the distribution of growth rates conditional on city size. Eeckhout (2004) who find that log-normal arises from Gibrat's law for all places, and Zipf's law can be a statistical phenomenon in the upper tail of the log-normal distribution. Eeckhout used 25,359 legally bounded places from U.S. census to represent all cities (incl. 1 person place) for the year 2000. Recently Berry and Okulica-Kozaryn (2012) using U.S. Economic Areas data from 1990 to 2010 find that Gibrat's law holds. They use simple cross-sectional OLS estimation: $logp_{it} = \alpha + \beta logp_{it-1} + \varepsilon_i$ (p_{it} represent population), if $\beta = 1.0$ then Gibrat's law holds. If the ε_i are iit normal the size distribution is Pareto (Malevergne *et al*., 2009, 2011). Results show that for the period 1990-2000, $\beta = 1.004$ and for 2000-2010 $\beta = 1.015$.

For other countries Gibrat's law can also be found. Clark and Stabler (1991) study Canada's 7 largest cities from 1975 to 1984 (from 1971 census metro area) and find that Gibrat's law cannot be rejected using hypothesis testing. Eaton and Eckstein (1997) use France and Japan's top 40 cities from 1876 to 1990 and find that there is no difference in the mean of growth rate between large and small cities. This indicates Gibrat's law holds and for France, the urban system displayed a significant pattern of parallel growth, where the average annual growth rate of urban population was unrelated to city scales; for Japan, city growth presented an approximately parallel growth model with a slightly divergent growth, where the average annual growth rate of urban population was positively, but statistically insignificantly correlated with city scales. Giesen and Sudekum (2010) find Gibrat's law not only holds at national level, but also tends to hold at regional level for Germany's 71 large cities [35] from 1975 to 1997 using non-parametric techniques as used by Ioannides and Overman (2003). Sharma (2003) for the first time applies Gibrat's law to developing countries. Sharma studies Indian cities from 1901 to 1991 and supports Gibrat's law by finding the parallel growth in the long run although with some deviations in the short run.

[35] With more than 100,000 inhabitants in 1997.

On the other hand, however, Gibrat's law is also rejected for some countries and periods. Firstly for the U.S. cities, Ioannides and Overman (2003) study the U.S. census data for metro areas from 1900 to 1990 using non-parametric methods [36] and find that Gibrat's law does not hold exactly for city growth process as both the mean and variance will vary with city size. Although there is constant mean and variance for city growth rate across all city sizes would lie within in the 5% confidence bands. Black and Henderson also find that Gibrat's law is rejected for any sample size for U.S. metropolitans from 1900 to 1990. As noted by Clark and Stabler (1991), testing for Gibrat's law is equivalent to testing for unit roots in the process of urban growth. Therefore, Blank and Henderson estimate the equation: $\ln(p_{it+1}) - \ln(p_{it}) = \alpha + \delta_t + \gamma \ln(p_{it}) + \varepsilon_{it}$ and test the null hypothesis that $\gamma = 0$ which implies Gibrat's law. Gamestani *et al.* (2007) study the U.S. south-eastern region cities from 1860 to 1990 (by decades) and also reject Gibrat's law. They find that city growth rate is correlated to its size, with small cities tending to have higher growth rates and large cities tending to have lower growth rates. Levy (2009) uses the same data as Eeckhout (2004) (the U.S. 25,359 legally bound places-U.S. census data for all cities incl. 1 person place) but finds that lognormal is rejected for large cities, which is against Gibrat's law. City size distribution is lognormal for the small and medium sized cities, but it follows a power law distribution for large cities. Glaeser *et al.* (2011) use U.S. county-level data (for the eastern region) from 1860 to 2000 and find that Gibrat's law

[36] Estimate a stochastic kernel- a three dimensional representation of the distribution of growth rates conditional on city size, using tsrf software.

seems to hold in the long-run but does not hold for some decades. The correlation between city growth and its initial size is close to 0 (-0.0034) between 1860 and 2000, but Gibrat's law is often broken in some decades. More recently, Gonzalez-Val *et al.* (2013) study the cities in U.S., Spain, Italy (each decade for 20th century), France, England and Japan (recent decades) and suggest that the parametric regressions for Gibrat's law may lead to biased results. Let S_{it} be the size (population) of a city i at time t, and $g$be its growth rate, then $S_{it} = (1 + g)S_{it-1}$. After taking logarithms and considering that the rate depends on the initial size, we can obtain the general expression of the growth equation: $lnS_{it} - lnS_{it-1} = \mu + \beta lnS_{it-1} + u_{it}$, where $\mu = \ln(1 + g)$ and u_{it} is an i.i.d. random variable representing the random shocks that influence a city's growth rate. Gibrat's law indicates$\beta = 0$, thus if the estimate of β is significant different from 0, we can reject Gibrat's law. If β is significant and positive, this indicates that city growth pattern is divergent. Then they estimate this equation for different sample sizes, namely 50, 100, 200, 500, 1000 cities and so on. Results are different depending on the sample size. The behaviour of the coefficient β is similar in all decades and for all the countries, for example, for the U.S. in the 1950-1960 period, from the small sample size to large sample size, the coefficient of β is observed non-monotonically conditional to sample size, beginning by accepting Gibrat's law, then a convergent behaviour to Gibrat's law (β is significant and negative), and finally a divergent behaviour (β is significant and positive). Therefore, they suggest that the appropriate methods to test for Gibrat's law are non-parametric methods (see Gabaix and Ioannides, 2004; Eeckhout, 2004). And a more precise definition of a city might be required.

Secondly, for other countries Gibrat's law is found to be rejected sometimes. Guerin-Pace (1995) studies French cities from 1831 to 1984 and finds that Gibrat's law does not hold as there are correlations between growth and size. Bosker *et al.* (2008) use 62 cities from West-Germany for 1925 to 1999 and find that Gibrat's law is rejected for about 75% of all cities by testing for the unit root. As mentioned above, if there is a unit root then city growth is under a stochastic process which will produce Gibrat's law, however, city growth is found to be trend stationary in Bosker *et al.* (2008)'s study. Dimou and Schaffar (2009) study the Balkan Peninsula for 1981, 1991 and 2001 and find that Gibrat's law does not hold. City growth follows a 'Hybrid' pattern, where city size filters the effects of common external shocks. Lastly, for the developing countries the studies are rare, Soo (2007) studies Malaysian cities from five censuses 1957, 1970, 1980, 1991, 2000 and the number of cities varies from 44 to 171. Soo finds that there is a negative correlation between city growth and size, i.e. smaller cities grow faster.

4.2.3 Empirical Evidence in China

The debate about city size distribution mentioned in the previous chapter (Chapter 3 Zipf's law) has already discussed some of the empirical evidence for Gibrat's law. As the study of Gibrat's law is rare in developing countries, for the case of China there are only a few studies testing for Gibrat's law. Generally speaking, Gibrat's law is rejected in China according to these studies but it seems to hold in the long run. Xu and Zhu (2008) study a panel data of 165 cities at prefecture-level and provincial-level cities from 1990 to 2000, and find that Gibrat's law is rejected using OLS, fixed effects and system GMM estimations (regress city growth on initial population level, ten-years growth over 1990-2000 and then five-year growth over two sub-periods, 1990-1995 and 1995-2000). In fact they study the urban growth determinants in China and find that the initial population level and GDP per capita of a city are two main factors that affect the urban growth in the 1990s. The initial population level has a persistently significant and negative sign on the urban growth rate which implies that small and medium-sized cities were growing faster than the large ones in 1990s, which is also consistent with the results of Xu and Zhu (2009). Xu and Zhu (2009) study the Chinese cites using non-agriculture populations for 1990-2000 (1990, 1993, 1995, 1998 and 2000). They test the convergence of city growth by estimating $\ln\left(\frac{N_{i,t+1}}{N_{i,t}}\right) = \alpha + \beta \ln\left(N_{i,t}\right) + \varepsilon_{it}$ if $\beta = 0$ then Gibrat's law holds. Results reject Gibrat's law with a negative β, but confirm that there is a convergence trend in urban growth in the 1990s. Wang and Zhu (2012) test Gibrat's law by time series Gini coefficients and panel unit root, using non-agriculture population of 48 Chinese cities, from 1949 to 2008. They find that Chinese city size distribution has shown an approximately parallel growth model in the long run, though it presented different patterns of growth in the short run.

In conclusion, with respect to Gibrat's law, it is generally rejected in Chinese cities, but there seems a trend that parallel urban growth will show in the long run (Wang and Zhu, 2012). However, there is also a voice of opposition of Anderson and Ge (2005); they study Chinese cities which are over 100,000 inhabitants from 1949 to 1999 (8 years respectively: 1949, 1961, 1970, 1985, 1990, 1994, 1999; the number of cities varies from 132 to 667) and find that city size distribution follows log-normal distribution using efficient OLS together with maximum likelihood and corresponding Pearson goodness of fit tests, which indicates that Gibrat's law holds.

4.3 METHODOLOGY

Gibrat's law indicates that population growth has typically been found to be essentially uncorrelated with initial population levels, both in the U.S., France and Japan (Glaeser, Scheinkman and Shleifer 1995; Eaton and Eckstein 1997; Glaeser and Shapiro 2003). The first step to test for Gibrat's law is to investigate the correlation coefficient between the growth rate and the initial population size, i.e. the correlation between the previous year's city growth rate and the population of current year, shown in Table 4.1. From 1880s to 2009, we find that the correlation coefficients are not significantly different from 0 for most of the years, indicating that the city growth rate is independent of its initial size, i.e. Gibrat's law holds.

Another commonly used method for testing for Gibrat's law is the parametric regression. Let $\boldsymbol{P_{i,t}}$ represent the population (size) of a city i at time$t$, following Soo (2014), Gibrat's law can be tested by estimating the following equation:

$$\ln\left(\boldsymbol{P_{i,t}}\right) = \boldsymbol{\mu_1} + \boldsymbol{\delta}_1 \, \boldsymbol{ln}\left(\boldsymbol{P_{i,t-1}}\right) + \boldsymbol{\theta_i} + \boldsymbol{\gamma_t} + \boldsymbol{\varepsilon_{i,t}} \quad (1)$$

where $\boldsymbol{\theta_i}$ is a set of cities fixed effects and $\boldsymbol{\gamma_t}$ is a set of year dummies, $\boldsymbol{\varepsilon_{i,t}}$ is the error term. If Gibrat's law holds, then δ =1 and the error term $\boldsymbol{\varepsilon_{i,t}}$ is i.i.d.. We denote it as model (1). Model (2) below follows Black and Henderson (2003) we estimate

$$\ln\left(\boldsymbol{P_{i,t}}\right) - \ln\left(\boldsymbol{P_{i,t-1}}\right) = \boldsymbol{\mu_2} + \boldsymbol{\delta}_2 \, \boldsymbol{ln}\left(\boldsymbol{P_{i,t-1}}\right) + \boldsymbol{\theta_i}' + \boldsymbol{\gamma_t}' + \boldsymbol{\varepsilon_{it}}' \quad (2)$$

The LHS is the change of population in logarithm and we aim to see how this change of population will be related to the initial size and also the lagged change that would be created as an instrumental variable automatically when using system GMM. For model (3) we add a quadratic term of initial size, to check the nonlinearity:

$$\ln\left(\boldsymbol{P_{i,t}}\right) - \ln\left(\boldsymbol{P_{i,t-1}}\right) = \boldsymbol{\mu_3} + \boldsymbol{\delta}_3 \, \boldsymbol{ln}\left(\boldsymbol{P_{i,t-1}}\right) + \rho_3 \, \boldsymbol{ln}\left(\boldsymbol{P_{i,t-1}}\right)^2 + \boldsymbol{\theta_i}'' + \boldsymbol{\gamma_t}'' + \boldsymbol{\varepsilon_{it}}'' \; (3)$$

In model (4) we calculate the current year growth rate by$(\boldsymbol{P_{i,t}} - \boldsymbol{P_{i,t-1}})/\boldsymbol{P_{i,t-1}}$, and regress on the initial population, which is in last year$\boldsymbol{P_{i,t-1}}$. Therefore, model (4):

$$\ln \frac{P_{i,t}-P_{i,t-1}}{P_{i,t-1}} = \mu_4 + \delta_4 \, ln\left(P_{i,t-1}\right) + \theta_i{}''' + \gamma_t{}''' + \varepsilon_{it}{}''' \qquad (4)$$

And model (5) we add the quadratic term of initial size:

$$\ln \frac{P_{i,t}-P_{i,t-1}}{P_{i,t-1}} = \mu_5 + \delta_5 \, ln\left(P_{i,t-1}\right) + \rho_5 \, ln\left(P_{i,t-1}\right)^2 + \theta_i{}'''' + \gamma_{t'}{}'''' + \varepsilon_{it}{}'''' \quad (5)$$

After obtaining the results for the whole sample, we then split the sample into prefecture-level cities and county-level cities, to see the fixed effects results. Finally, we divide the whole sample into four economic regions, East, Midland, West and North-East, to run these 5 models as above. This is for the first time to investigate the distribution and growth pattern of Chinese regional cities. Besides, we also add year dummy of 1984, as this is nearly the start of "Chinese Economic Reform" and the "One Child per Family" policy" [37]as well.

In addition, Resende (2004) explores the dynamic implications of Gibrat's law in terms of a unit root for the log of city size. Gibrat's law indicates that, regardless of initial size, all cities grow randomly with the same expected rate and same variance. If there is a unit root in city size, then the process of city size is not stationary but stochastic, i.e. cities grow stochastically, which indicates Gibrat's law might hold. Actually, Eq. (11) is equivalent to a panel unit root test:

[37] 'Chinese Economic Reform and 'One Child Policy' were launched in 1979 and were starting to be effective at mid-1980s, like 1984.

$$\ln\left(P_{i,t}\right) - \ln\left(P_{i,t-1}\right) = \mu_2 + (b-1)\, ln\left(P_{i,t-1}\right) + \theta_i{}' + \gamma_t{}' + \varepsilon_i{}' \quad (11)'$$

to test whether $|b|$<1, if $|b|$<1 then there is a convergent reversion to steady state of log population. Otherwise, $|b| \leq 1$ the log population would follow a stochastic process, indicating Gibrat's law holds. We use Fisher- Augmented Dickey-Fuller test and Fisher-Phillips-Perron test to test the panel unit root of logarithm of city size. The null hypothesis is that the variable contains a unit root, and the alternative is that the variable was generated by a stationary process.

4.4 RESULTS

As mentioned previously, researchers commonly agree that Gibrat's law is one of the theoretical explanations of Zipf's law, i.e. Gibrat's law can deliver Zipf's law. If cities grow under the mode of Gibrat's law, i.e. grow stochastically with a growth process that is homogeneous with a common mean and variance, and then in the steady state the growth process will produce a city size distribution following Zipf's law (Gabaxi, 1999). Therefore, in this chapter we test for Gibrat's law- a model of city growth arguing that cities grow homogeneously with a common mean and variance. The derivation of Gibrat's law indicates that city growth rate is independent of its initial size. Originally, Gibrat's law known also as the proportionate growth, i.e. increment of population is proportionate to its initial size, which can be presented as (Steindl, 1965):

$$x_t - x_{t-1} = \varepsilon_t x_{t-1}$$

where x_t is the population size of a city at time t and ε_t is the proportional rate of growth between period (t-1) and t. Then

$$\frac{x_t - x_{t-1}}{x_{t-1}} = \varepsilon_t$$

indicates that city growth rate is independent of its initial size. To conclude, the population change is proportionate to initial size but the growth rate is not. Therefore, first of all we test for the correlation between subsequent city growth rate and its initial size.

4.4.1 Correlation between Population Growth Rate and Initial Size

Gibrat's law states that within a country the population growth rate of a city is independent of its initial size, i.e. there is no correlation between population growth and initial population. Therefore, the first direct way to test for Gibrat's law is to investigate the correlations. Table 4.1 below reports the correlation coefficient between subsequent population growth rate and the initial city size (current year population) for available cities for each year. If Gibrat's law holds we would expect the correlation coefficient to be 0. Then we also investigate the subsamples for prefecture-level cities, county-level cities; East cities, Midland cities, West cities and North-East cities; 4 subgroups according to cities' age. The most unique contribution of this paper is that we divide the whole sample into 4 groups of historical cities.

(1) Whole sample.

First of all, for the whole sample, correlation coefficients are reported in column 2 Table 4.1 and the evolution of the coefficient is shown in Figure 4.1 below. Two features can be observed. Firstly, the correlation coefficient between current year city growth rate and initial size (previous year population) is not significantly different from 0 for most of the years (5 years exception out of 28 years), which indicates that city growth rate is independent of its initial size, i.e. the prediction of Gibrat's law. Moreover, the values of the correlation coefficients are quite small around -0.16 to 0.056, and have a trend close to 0, showing that Gibrat's law approximately holds during 1980s, 1990s and 2000s and there is a trend that city growth is closing to Gibrat's law (the coefficients are closing to 0). The value of correlation coefficients are almost all negative for the whole period from 1981 to 2009 (except for 3 years in 1995, 1998 and 2009) which indicates that initially large city's growth rate would be relatively small in the next year; large cities were not growing faster than smaller ones for most of the years in the last three decades (after the 'Economic Reform' launched). However, this effect is quite small due to the coefficient being not significantly different from 0 and having small values. This is consistent with Gabaix (1999) who argues that Zipf's law comes from the Gibrat's law mode of urban growth. Recall that for the whole sample Zipf's law is found during the middle of the urban growth process in the end of the 1980s and early 1990s, this is because Chinese cities are growing in a manner that is consistent with Gibrat's law. But unlike Gabaix (1999) who argues that the steady state of Gibrat's law-"mode" of growth is Zipf's law, Chinese city size distribution is consistent with Zipf's law for a few years and then becomes more even.

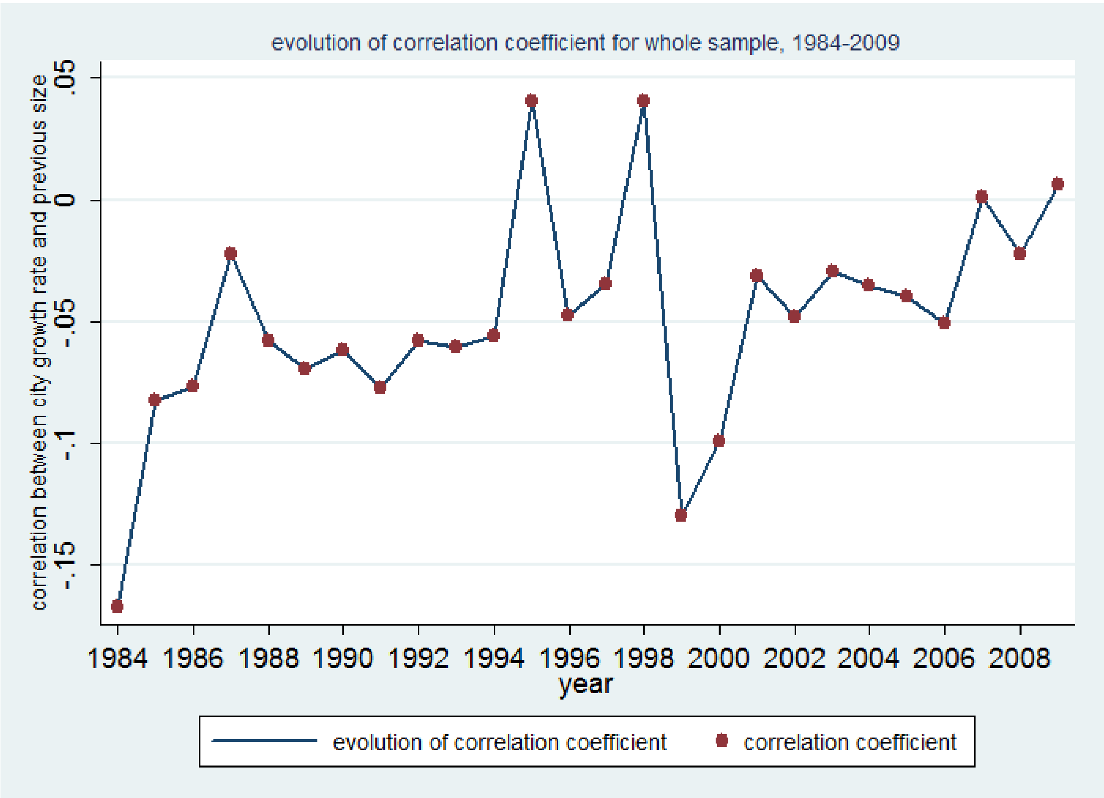


Figure 4.1: Evolution of correlation coefficient between city growth rate and initial size - whole sample

(2) Prefecture-level cities and county-level sample.

As showed in Figure 4.2 below and the correlation coefficient in Table 4.1, for prefecture-level cities, the evolution of the correlation coefficient has quite a similar shape as the whole sample and is not significantly different from 0 (4 years exception out of 28 years), which lead to similar results as the whole sample i.e. that Gibrat's law approximately holds. This might also reveal that the main body of urban growth is prefecture-level cities due to their administrative position and long history of population agglomeration. In contrast, the county-level cities do not support Gibrat's law. The correlation coefficient, shown in Table 4.1, is significantly different from 0 for almost half of the years studied (12 years out of 28 years), i.e. city growth rate is not independent of initial size, Gibrat's law does not hold. This is also evident from the value of the correlation coefficient, which is not close to 0 and fluctuates a lot compared to prefecture-level cities' sample. To conclude, the results for Gibrat's law is consistent with the results for Zipf's law as analysed before. Prefecture-level cities tend to grow under Gibrat's law-mode of growth, thus deliver Zipf's law in the middle of the growth process for some years (1958-1983), then prefecture-level city size distribution become more even than Zipf's prediction. In contrast, county-level cities do not show Gibrat's law-mode of growth and thus do not show Zipf's law for the whole period from 1984 to 2009. But they have a trend growing closer to Zipf's distribution, although have not yet reached Zipf's prediction (Pareto exponent grows but stays relatively stable at 0.8 in previous chapter).

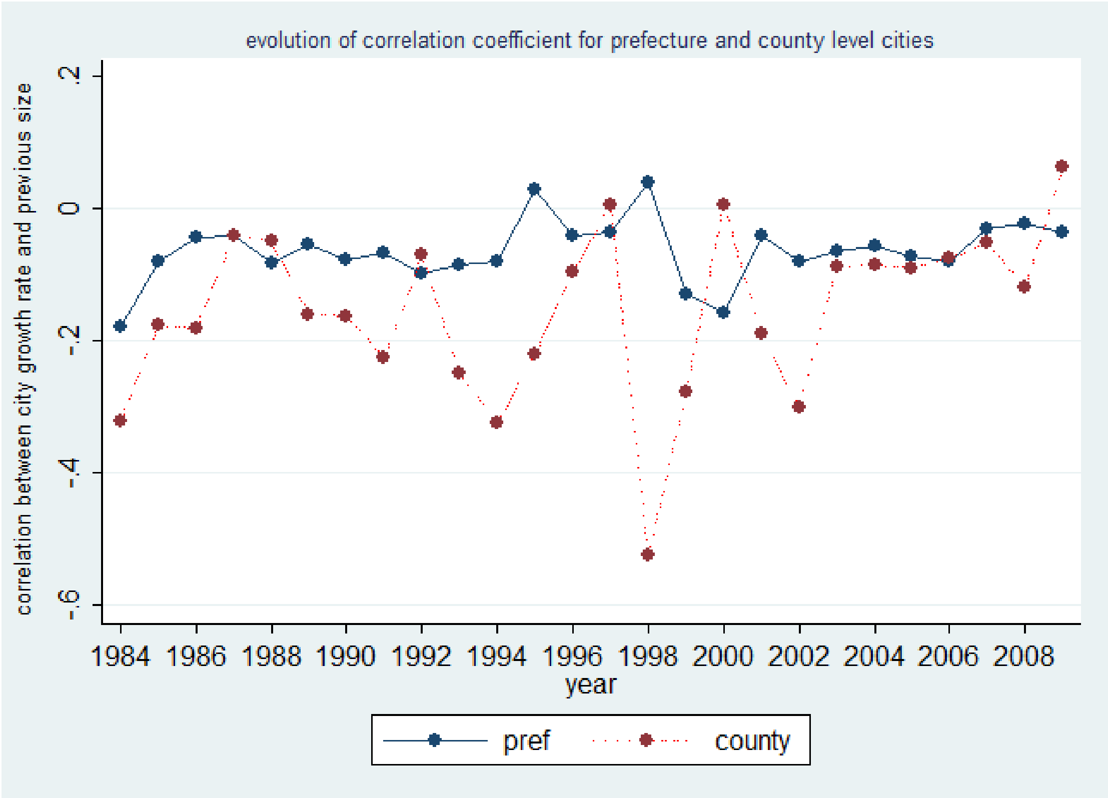


Figure 4.2: Evolution of correlation coefficient between city growth rate and initial size - prefecture and county-level cities

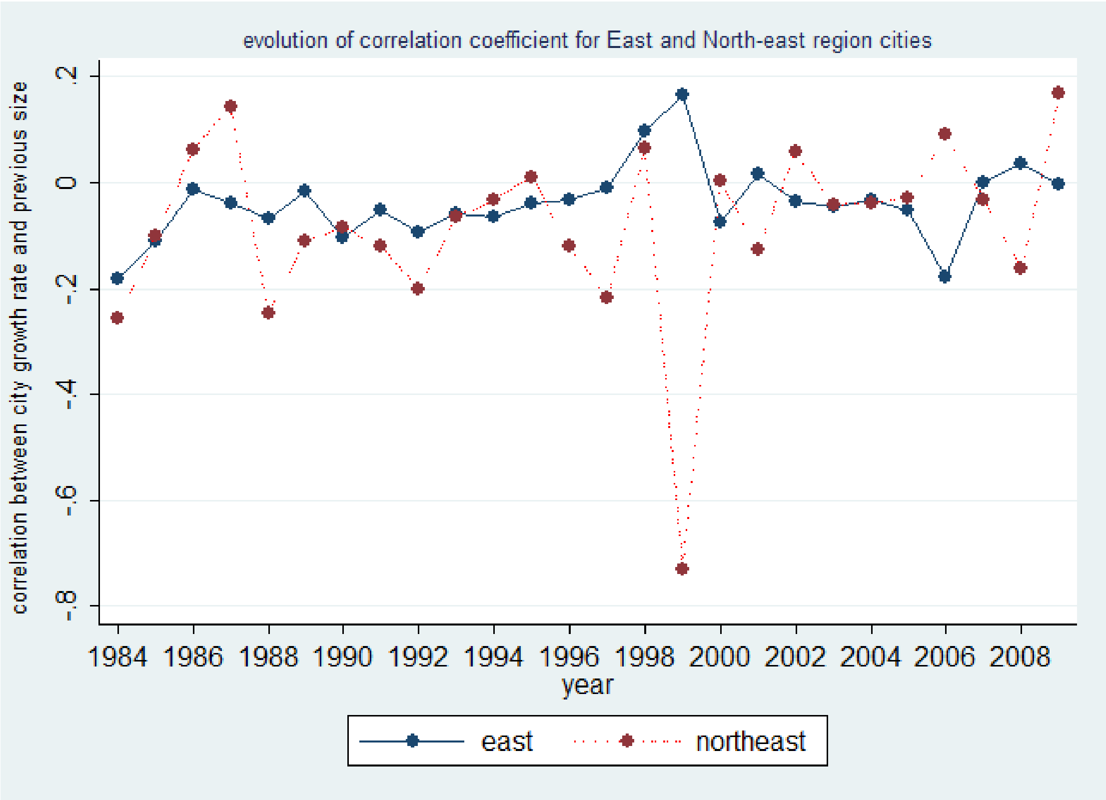
(3) Regional sample.

We now consider the urban growth pattern of the regional cities. Figure 4.3 below shows the evolution of the correlation coefficient for four economic regions. It would appear that Gibrat's law approximately holds for the East cities, because firstly from Table 4.1 below, we can see that most of the correlation coefficient is not statistically significantly different from 0 (2 years exception out of 28 years). Secondly from the evolution of coefficient graph below, the east cities' correlation coefficients are relatively stable and the values are quite close to 0, which indicates city growth rate is independent of initial size, i.e. Gibrat's law approximately holds. Then one can expect

Zipf's law from Gibrat's law, which did happen in East region as mentioned in previous chapter that Zipf's law is found in East cities. It is worth noting that most of the correlation coefficients are smaller than 0 for most of the time, which means that large cities were growing not so fast as smaller cities in East region China, then city size distribution grew evenly as Zipf's law predicts.

In contrast, for other regions Gibrat's law does not seem to hold, as the correlation coefficients for the other regions fluctuate quite widely. Specifically, for the North-East cities which is close to the East cities geographically, the correlation coefficient is smaller than 0 but not far away from 0 for most of the time (-0.25 to 0.14, expect for one year 1999 the coefficient is -0.73), which shows that there might be a trend that city size distribution becomes more even as large cities are growing a little bit less faster than smaller ones. For the Midland and West regions, the correlation coefficient is typically smaller than 0 for most of the years with a weak trend closing to 0 over time, which still shows a convergent growth process that large cities might grow not as fast as smaller ones. To be noted that for the year 1999, large cities grew relatively slower in the East region but faster in the Midland region and West region (1998) than smaller ones, which is consistent with the 'Western and Midland

Development' policy in the end of 1990s.[38]



[38] 'Economic Reform' was launched in 1879 giving the priority of development for East region. After two decades development, East region is much more developed than Midland and West regions, therefore, Chinese government launched another policy about 'Western and Midland Development'. The North-East region has long been the centre of heavy-industry, prosperous in late 1980s (when large cities grow faster as showed in evolution graph, but declined a bit in 1990s (small cities grow faster as showed in evolution graph), then Policy of 'The revitalization of the old industrial base of the Northeast' launched science 2003.

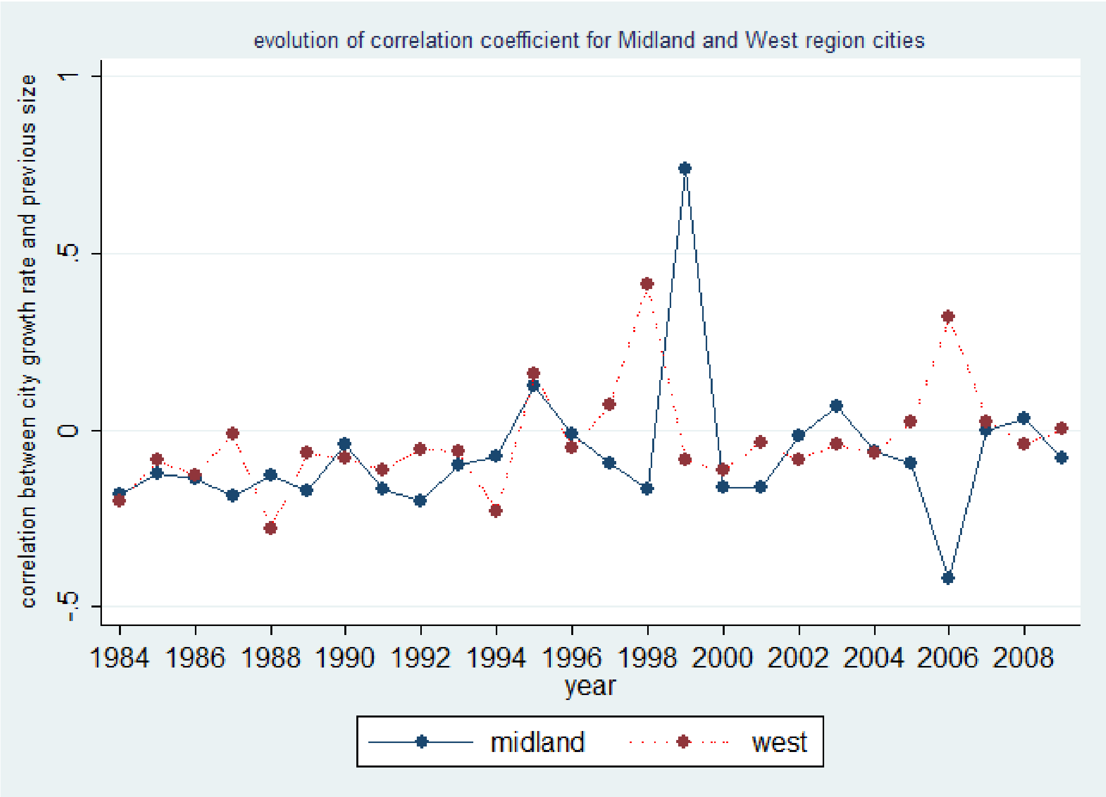


Figure 4.3: Evolution of correlation coefficient between city growth rate and initial size - regional cities

(4) **Historical sample.**

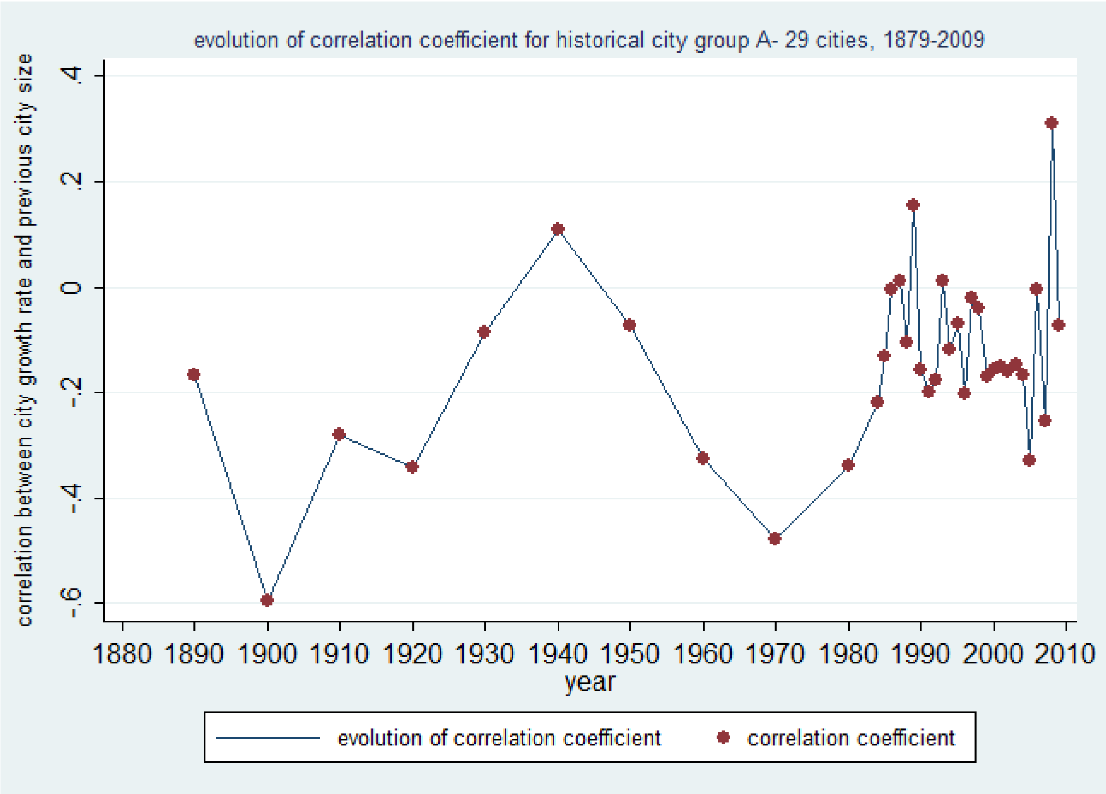
For the first time in the literature, Chinese cities are divided into 4 historical city groups according to their age. The evolution of the correlation coefficients for these four subgroups is shown in Figure 4.4 below. Firstly, on the whole, as shown in the four graphs, most of the correlation coefficients are negative in each graph, which indicates that the growth rate for these cities is negative relates to their size, i.e. large cities tend to grow slower than small ones in each sample. However there seems to be a trend for every graph that the coefficients are becoming closer to 0 over time, i.e. the city growth is closing to Gibrat's law over time. In addition, Table 4.1 below shows the value of the correlation coefficients and the testing for whether the coefficients are significantly different from 0. Most of the coefficients are not significantly different from 0 (1 year exception in group A, 2 years in group B, 1 year in group C and 1 year in group D), i.e. previous year's city growth rate is independent of current city size, which indicates Gibrat's law holds.

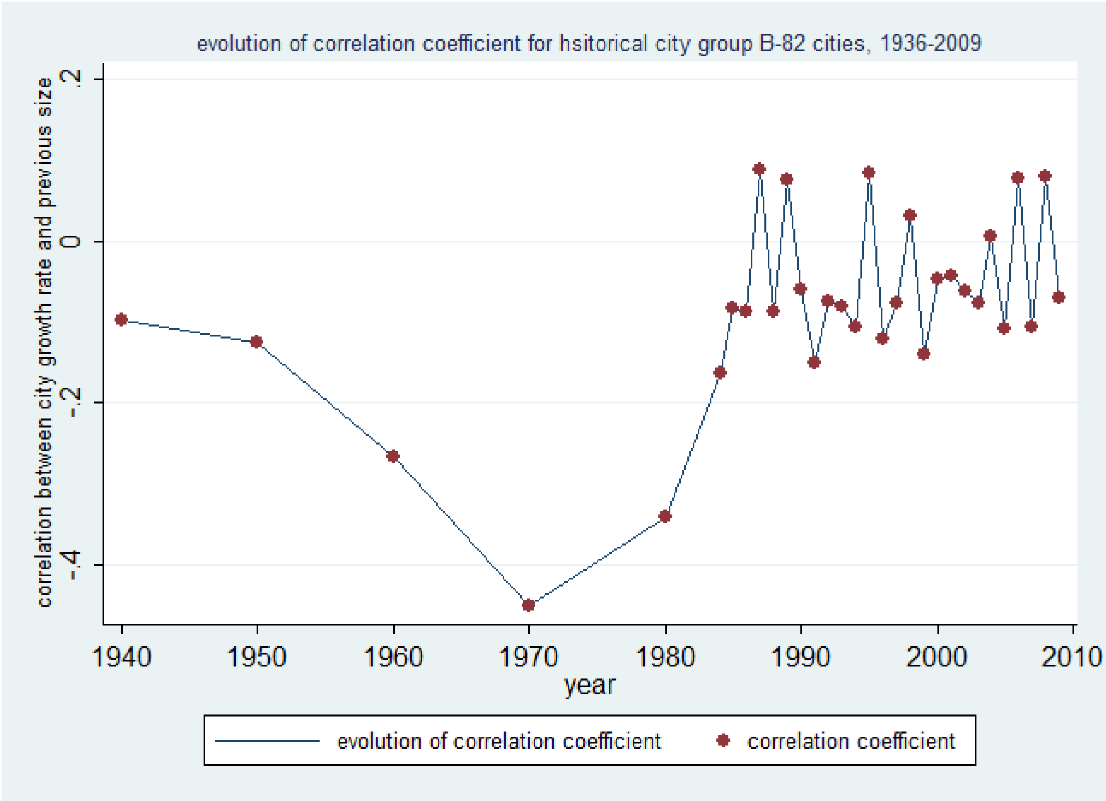
Specifically, the first graph in Figure 4.4 shows the group A- 29 historical cities from 1879 to 2009 (balanced panel). These cities exist from the end of the 'Qing' dynasty until now and experience a lot of shocks including the WWI, WWII, national war, stagnant period, 'One Child policy' and 'Economic reform'. The graph shows a trend that the correlation coefficient is tending towards 0 over this long period, although the coefficients are negative for most of the time which indicates that the growth rate for these long-history cities is negatively related to their size, i.e. large cities tend to grow slower than small ones in this sample. But the coefficients are not significantly different from 0.

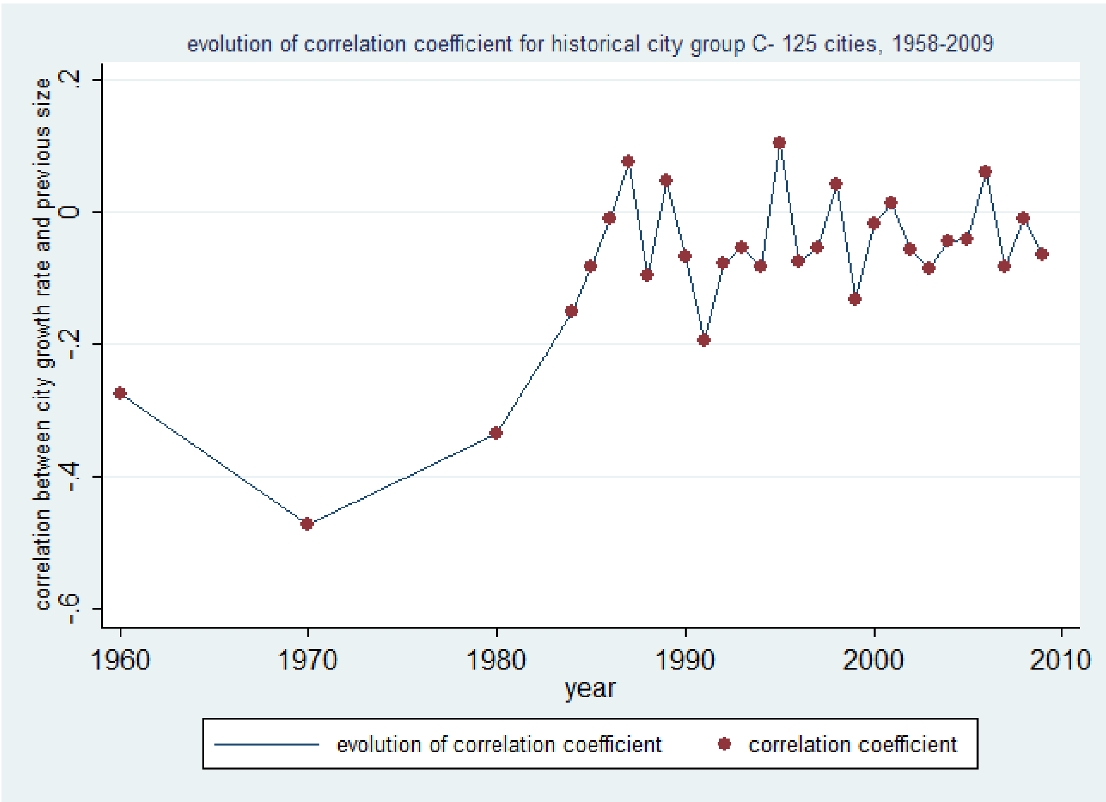
The second graph in Figure 4.4 is the group B- 82 historical cities from 1936 to 2009. This sample starts from the early stage of National government ruled by KMT until now, which experienced the same as group A except for the WWI. Similarly, most of the coefficients are negative, but after 1980 the coefficients are around 0. The third graph in Figure 4.4 shows the group C- 125 cities from 1958 to 2009, which shows the cities exist from the first decade of the PRC China (ruled by communist party) until now. These cities were growing under a relatively more stable political and economic environment than previous periods, no wars, but they did experience the stagnant period. The correlation coefficients are basically the same with group B, having a trend towards to 0 and after 1980 the coefficients are around 0 (fluctuating from -0.2 to 0.2).

The last group in Figure 4.4 shows the group D- 294 cities from 1984 to 2009. Cities in this group grew relatively stochastically under the stable political and economic environment after the 'Economic reform'. Thus, the growth pattern for these cities tends to be close to Gibrat's law. As shown in the graph, the coefficients are much closer to 0, fluctuating from -0.15 to 0.06; and the value of the coefficients are not significantly different from 0 in Table 4.1. These indicate Gibrat's law holds for these cities.

To conclude, these four groups of historical cities show the evolution of city growth towards a Gibrat's-mode of urban growth. Before the early 1980s, large cities seem to grow a bit slower than smaller ones as correlation coefficients are constantly negative (except for 1940); since the 'Economic Reform' pursuing 'development come first, equality second' which favours large cities, large cities begin to grow faster than smaller ones (correlation coefficient is positive sometimes) and then the correlation coefficients are fluctuating and becoming closer to 0 over time, which indicates that the growth rate becomes independent of initial size gradually. In addition, the values of the correlation coefficients are consistently significantly not different from 0, which also indicates the city growth rate is independent of its initial size, i.e. Gibrat's law.




evolution of correlation coefficient for hsitorical city group B-82 cities, 1936-2009
correlation between city growth rate and previous size
-.4
-.2
0
.2
1940
1950
1960
1970
1980
1990
2000
2010
year
evolution of correlation coefficient
correlation coefficient


evolution of correlation coefficient for historical city group C- 125 cities, 1958-2009
correlation between city growth rate and previous size
-.6
-.4
-.2
0
.2
1960
1970
1980
1990
2000
2010
year
evolution of correlation coefficient
correlation coefficient

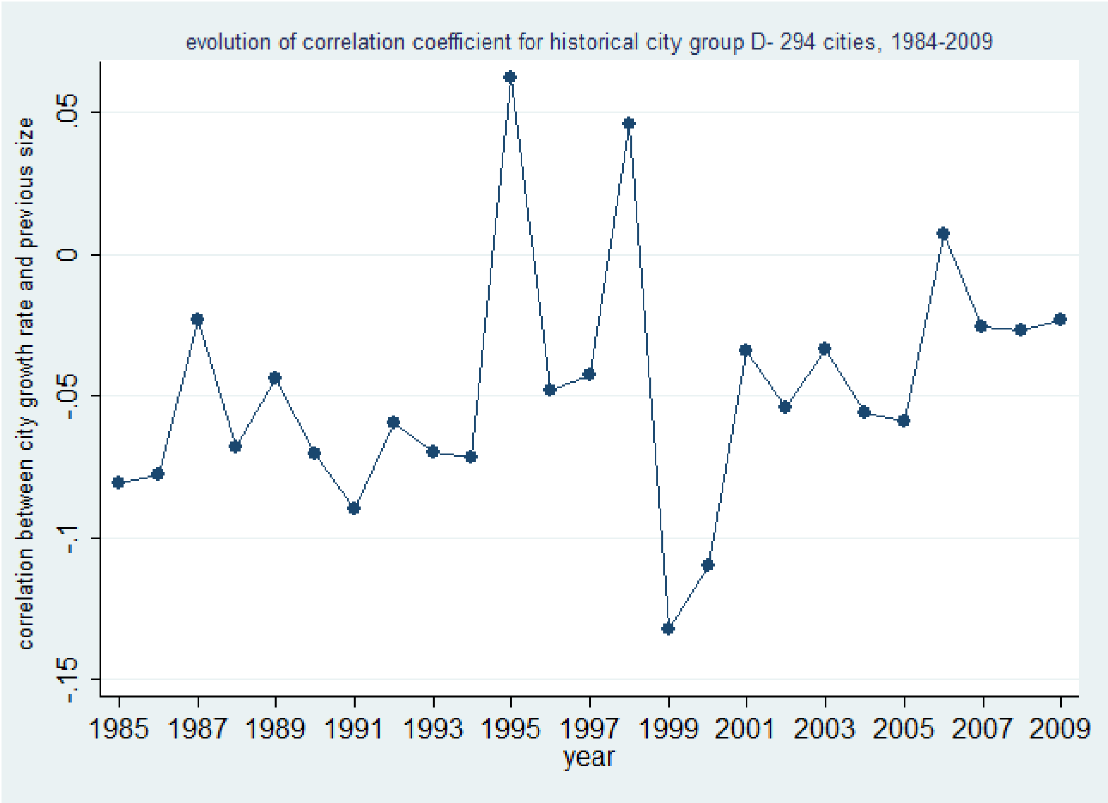


Figure 4.4: Evolution of correlation coefficient between city growth rate and initial size - historical cities

In conclusion, from the correlation coefficients between city growth and initial city size for each year, obviously we can find that in most cases Gibrat's law approximately holds for Chinese cities over 1879-2009, as most of the coefficients of correlation are not significantly different from 0 and the values are not far away from 0. In addition, there is a trend that the coefficients of correlation are becoming closer to 0 over time, which also indicates that there is a trend for city growth pattern towards the Gibrat's-mode of growth, especially in East region and 4 groups of historical cities.

Table 4.1: Correlations between population growth and the initial city size (initial population)

	Whole sample	prefectural-level cities	county-level cities	East cities	Midland cities	West cities	North-East cities	29 cities 1879-2009	82 cities 1936-2009	125 cities 1958-2009	294 cities 1984-2009
1880s											
1890s								-0.1675			
								(0.5353)			
1900s								-0.5942**			
								(0.0195)			
1910s								-0.2816			
								(0.3513)			
1920s								-0.3425			
								(0.1784)			
1930s								-0.0862			
								(0.7423)			
1940s								0.1091	-0.0973		
								(0.5958)	(0.3844)		
1950s								-0.0717	-0.1251		
								(0.7278)	(0.2627)		
1960s								-0.3261	-0.2672*	-0.2732	
								(0.1491)	(0.0664)	(0.0500)	
1970s								-0.4791	-0.4521***	-0.4727***	
								(0.0241)	(0.0013)	(0.0004)	
1980s								-0.3405	-0.3420**	-0.3352	
								(0.1210)	(0.0173)	(0.0151)	
1984	-0.1675**	-0.1775*	-0.3207	-0.1811	-0.1822	-0.2028	-0.2557	-0.2181	-0.1638	-0.1485	
	(0.0227)	(0.0218)	(0.1944)	(0.1738)	(0.2006)	(0.1623)	(0.1980)	(0.2649)	(0.1575)	(0.1132)	
1985	-0.0823	-0.0806	-0.1759	-0.1090	-0.1253	-0.0845	-0.1012	-0.1297	-0.0834	-0.0817	-0.0811
	(0.1593)	(0.2265)	(0.1545)	(0.3063)	(0.2619)	(0.4339)	(0.5691)	(0.5107)	(0.4591)	(0.3689)	(0.1656)
1986	-0.0768	-0.0443	-0.1810	-0.0141	-0.1379	-0.1287	0.0615	-0.0059	-0.0864		-0.0780
	(0.1832)	(0.5050)	(0.1254)	(0.8995)	(0.2197)	(0.1996)	(0.7176)	(0.9763)	(0.4492)	-0.0103	(0.1921)

1987	-0.0224	-0.0403	-0.0404	-0.0395	-0.1839*	-0.0102	0.1444	0.0122	0.0879	0.0772	-0.0235
	(0.6900)	(0.5429)	(0.7067)	(0.7120)	(0.0862)	(0.9186)	(0.3870)	(0.9508)	(0.4409)	(0.4001)	(0.6956)
1988	-0.0580	-0.0827	-0.0489	-0.0673	-0.1275	-0.2768***	-0.2469*	-0.1056	-0.0867	-0.0957	-0.0683
	(0.2596)	(0.2016)	(0.5673)	(0.4632)	(0.2038)	(0.0041)	(0.0807)	(0.5858)	(0.4415)	(0.2925)	(0.2464)
1989	-0.0694	-0.0530	-0.1596**	-0.0166	-0.1740*	-0.0661	-0.1091	0.1559	0.0760	0.0478	-0.0440
	(0.1503)	(0.4039)	(0.0319)	(0.8435)	(0.0678)	(0.4811)	(0.4068)	(0.4283)	(0.5027)	(0.6007)	(0.4574)
1990	-0.0615	-0.0762	-0.1617**	-0.1032	-0.0419	-0.0782	-0.0841	-0.1555	-0.0586	-0.0655	-0.0707
	(0.1946)	(0.2297)	(0.0232)	(0.2015)	(0.6580)	(0.4041)	(0.5158)	(0.4296)	(0.6059)	(0.4732)	(0.2320)
1991	-0.0776*	-0.0657	-0.2250***	-0.0516	-0.1667*	-0.1123	-0.1214	-0.1982	-0.1510	-0.1950	-0.0899
	(0.0943)	(0.2980)	(0.0009)	(0.5103)	(0.0700)	(0.2239)	(0.3431)	(0.3027)	(0.1783)	(0.0307)	(0.1259)
1992	-0.0582	-0.0983	-0.0692	-0.0927	-0.1992**	-0.0524	-0.2012	-0.1774	-0.0752	-0.0772	-0.0597
	(0.2050)	(0.1181)	(0.3045)	(0.2249)	(0.0299)	(0.5701)	(0.1109)	(0.3572)	(0.5047)	(0.3963)	(0.3102)
1993	-0.0602	-0.0841	-0.2489*	-0.0576	-0.0968	-0.0611	-0.0641	0.0125	-0.0800	-0.0539	-0.0697
	(0.1720)	(0.1787)	(0.0001)	(0.4239)	(0.2849)	(0.4970)	(0.5954)	(0.9488)	(0.4776)	(0.5535)	(0.2356)
1994	-0.0560	-0.0804	-0.3242***	-0.0652	-0.0742	-0.2313***	-0.0324	-0.1190	-0.1066	-0.0823	-0.0720
	(0.1836)	(0.1964)	(0.0000)	(0.3311)	(0.4070)	(0.0067)	(0.7766)	(0.5385)	(0.3436)	(0.3654)	(0.2219)
1995	0.0409	0.0305	-0.2194***	-0.0394	0.1271	0.1598*	0.0111	-0.0696	0.0845	0.1047	0.0620
	(0.3107)	(0.6248)	(0.0000)	(0.5266)	(0.1434)	(0.0566)	(0.9219)	(0.7196)	(0.4504)	(0.2470)	(0.2941)
1996	-0.0475	-0.0393	-0.0957*	-0.0318	-0.0116	-0.0516	-0.1216	-0.2020	-0.1210	-0.0739	-0.0480
	(0.2311)	(0.5278)	(0.0634)	(0.6034)	(0.8933)	(0.5338)	(0.2765)	(0.2933)	(0.2790)	(0.4149)	(0.4169)
1997	-0.0345	-0.0355	0.0067	-0.0105	-0.0956	0.0712	-0.2172	-0.0196	-0.0758	-0.0542	-0.0426
	(0.6086)	(0.6034)	(0.9887)	(0.9235)	(0.4833)	(0.6270)	(0.2325)	(0.9211)	(0.5010)	(0.5654)	(0.5501)
1998	0.0405	0.0412	-0.5232	0.0973	-0.1683	0.4122***	0.0640	-0.0394	0.0318	0.0420	0.0457
	(0.5478)	(0.5473)	(0.2282)	(0.3727)	(0.2149)	(0.0033)	(0.7280)	(0.8420)	(0.7783)	(0.6560)	(0.5220)
1999	-0.1300*	-0.1292*	-0.2781	0.1648	0.7372***	-0.0838	-0.7288***	-0.1687	-0.1390	-0.1327	-0.1322
	(0.0500)	(0.0552)	(0.5459)	(0.1250)	(0.0000)	(0.5589)	(0.0000)	(0.3907)	(0.2157)	(0.1576)	(0.0601)
2000	-0.0994**	-0.1561**	0.0061	-0.0740	-0.1616*	-0.1141	0.0036	-0.1573	-0.0459	-0.0175	-0.1102

	(0.0116)	(0.0127)	(0.9046)	(0.2223)	(0.0573)	(0.1719)	(0.9737)	(0.4152)	(0.6822)	(0.8479)	(0.0671)
2001	-0.0314	-0.0415	-0.1885***	0.0175	-0.1622*	-0.0351	-0.1261	-0.1493	-0.0434	0.0131	-0.0340
	(0.4241)	(0.5068)	(0.0002)	(0.7730)	(0.0529)	(0.6733)	(0.2475)	(0.4395)	(0.6988)	(0.8855)	(0.5735)
2002	-0.0485	-0.0795	-0.2990***	-0.0365	-0.0176	-0.0825	0.0586	-0.1615	-0.0615	-0.0556	-0.0543
	(0.2180)	(0.1944)	(0.0000)	(0.5546)	(0.8341)	(0.3106)	(0.5923)	(0.4028)	(0.5829)	(0.5414)	(0.3646)
2003	-0.0294	-0.0636	-0.0881*	-0.0460	0.0657	-0.0413	-0.0418	-0.1472	-0.0757	-0.0848	-0.0333
	(0.4539)	(0.2953)	(0.0880)	(0.4586)	(0.4354)	(0.6065)	(0.7023)	(0.4460)	(0.4990)	(0.3511)	(0.5785)
2004	-0.0351	-0.0562	-0.0842	-0.0332	-0.0601	-0.0661	-0.0380	-0.1653	0.0071	-0.0430	-0.0560
	(0.3725)	(0.3514)	(0.1058)	(0.5926)	(0.4758)	(0.4077)	(0.7329)	(0.3915)	(0.9497)	(0.6383)	(0.3510)
2005	-0.0399	-0.0714	-0.0899*	-0.0519	-0.0952	0.0209	-0.0303	-0.3281	-0.1075	-0.0399	-0.0592
	(0.3094)	(0.2356)	(0.0829)	(0.4033)	(0.2598)	(0.7897)	(0.7854)	(0.0823)	(0.3364)	(0.6624)	(0.3247)
2006	-0.0511	-0.0805	-0.0744	-0.1795***	-0.4180***	0.3183***	0.0898	-0.0061	0.0776	0.0619	0.0071
	(0.1910)	(0.1809)	(0.1482)	(0.0036)	(0.0000)	(0.0000)	(0.4108)	(0.9750)	(0.4883)	(0.4983)	(0.9058)
2007	0.0010	-0.0297	-0.0509	-0.0004	-0.0005	0.0253	-0.0324	-0.2530	-0.1059	-0.0811	-0.0259
	(0.9788)	(0.6214)	(0.3271)	(0.9946)	(0.9956)	(0.7489)	(0.7701)	(0.1855)	(0.3436)	(0.3747)	(0.6667)
2008	-0.0221	-0.0223	-0.1180*	0.0346	0.0319	-0.0410	-0.1623	0.3092	0.0792	-0.0087	-0.0271
	(0.5722)	(0.7110)	(0.0225)	(0.5775)	(0.7040)	(0.6045)	(0.1378)	(0.1027)	(0.4795)	(0.9242)	(0.6520)
2009	0.0063	-0.0341	0.0638	-0.0019	-0.0794	0.0031	0.1690	-0.0722	-0.0702	-0.0629	-0.0234
	(0.8722)	(0.5696)	(0.2191)	(0.9761)	(0.3439)	(0.9691)	(0.1221)	(0.7096)	(0.5308)	(0.4913)	(0.6966)

Significance of correlation coefficient is showed as *** p<0.01, ** p<0.05, * p<0.1. p-value is showed in parentheses.

Note: We use 1879, 1890, 1900, average of 1911 and 1918, average of 1920 and 1926, average of 1930, 1936 and 1939, 1948, average of 1953 and 1958, average of 1961 and 1965, average of 1970 and 1975 to represent 1880s, 1890s, 1900s, 1910s, 1920s, 1930s, 1940s, 1950s, 1960s, 1970s.

1900's urban population data is only available for a few cities due to the WWI; therefore the population of cities and relative growth rate takes the average value of 1890s and 1910s.

County-level cities mostly established from 1980s.

4.4.2 City Growth Regression

After showing the correlation coefficients between city growth and initial city size for each year, obviously we can find that in most cases Gibrat's law approximately holds for Chinese cities over 1879-2009. We can also test for Gibrat's law by estimating the following Eq. (1) - Eq. (5), using unbalanced panel data from 1879 to 2009.

As shown in section 4.3- methodology part, we perform 5 models Eq. (1) - Eq. (4) to test for Gibrat's law using system GMM, two step. Recall these 5 equations stated in section3:

$$\boldsymbol{ln\left(P_{i,t}\right) = \mu_1 + \delta_1 \, ln\left(P_{i,t-1}\right) + \theta_i + \gamma_t + \varepsilon_i} \qquad (1)$$

$$\boldsymbol{ln\left(P_{i,t}\right) - ln\left(P_{i,t-1}\right) = \mu_2 + \delta_2 \, ln\left(P_{i,t-1}\right) + \theta_i{}' + \gamma_t{}' + \varepsilon_i{}'} \qquad (2)$$

$$\boldsymbol{ln\left(P_{i,t}\right) - ln\left(P_{i,t-1}\right) = \mu_3 + \delta_3 \, ln\left(P_{i,t-1}\right) + \rho_3 \, ln\left(P_{i,t-1}\right)^2 + \theta_i{}'' + \gamma_t{}'' + \varepsilon_i{}''} \quad (3)$$

If Gibrat's law holds, according to the context of Gibrat's law, then we will expect that $\boldsymbol{\delta}_1 = 1, or\ \boldsymbol{\delta}_2 = 0, or\ \boldsymbol{\delta}_3 = 0\ and\ \rho_3 = 0$ In Table 4.2 below we can find none of the result for any regressions strictly support Gibrat's law in Chinese cities; just some of the subgroups may be more close to Gibrat's law in Table 4.3. First of all, for the whole sample estimation, the result for $\boldsymbol{\delta}_1$ is 0.83 and significantly, which indicates that Gibrat's law is rejected for whole Chinese cities. Then for model 2, $\boldsymbol{\delta}_2$ is significantly different from 0 at 1% significance level, which implies the current population change is significantly related to initial size, i.e. Gibrat's law does not hold for whole Chinese cities. And for model 3, we add the quadratic term of initial city size, change the dependent variable to city growth rate and add the quadratic term in growth rate equation, still not support Gibrat's law for Chinese cities.

Table 4.2: Fixed effect estimation for whole sample

	(1)	(2)	(3)
	m1	m2	m3
VARIABLES	lnpop	lngrowth	lngrowth
L.lnpop	0.830***	-0.170***	-0.377**
	(0.0168)	(0.0168)	(0.158)
L_1lnpop2			0.0166
			(0.0120)
Constant	1.111***	1.111***	1.744***
	(0.107)	(0.107)	(0.513)
Observations	12,813	12,813	12,813
R-squared	0.761	0.118	0.122
Number of code	702	702	702

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Then we investigate whether Gibrat's law holds in prefecture-level cities or county-level cities; East or Midland or West or North-East cities; and whether it holds in four historical city groups. Firstly, From Table 4.3- Panel A below, we find that the result for Prefecture-level cities and county-level cities are quite similar to the whole sample- Gibrat's law does not hold strictly. From Table 4.3- Panel B in the four economic regions, we find only for East cities, $\boldsymbol{\delta_1}$ is closer to 1 (0.875). While for the four historical city groups in Table 4.3-Panel C we find that for the cities with older histories the growth processes is more close to Gibrat's law, as for $\boldsymbol{\delta_1}$ is increasing from modern cities group (294 cities, 1984-2009) to historical cities group (29 cities, 1879-2009). And the estimated coefficients for $\boldsymbol{\delta}_1$ are the closest to 1 compared to the other subsamples (around 0.9). Except for the 294 cities (1984-2009) group, other historical cities' $\boldsymbol{\delta}_1$ are all greater than 0.9, which is closer to Gibrat's law. This implies that in China the growth process of historical cities are closer to Gibrat's law, whose underlying process is similar to a stochastic movement thus the distribution is closer to lognormal.

Table 4.3_ Panel A: fixed effect for subsamples- Prefecture-level cities

	Prefecture-level cities			County-level cities		
	m1	m2	m3	m1	m2	m3
VARIABLES	lnpop	lngrowth	lngrowth	lnpop	lngrowth	lngrowth
L.lnpop	0.832***	-0.168***	-0.820***	0.815***	-0.185***	0.276**
	(0.0184)	(0.0184)	(0.167)	(0.0423)	(0.0423)	(0.125)
L_1lnpop2			0.0508***			-0.0452***
			(0.0123)			(0.0135)
Constant	1.135***	1.135***	3.200***	1.158***	1.158***	0.0663
	(0.121)	(0.121)	(0.560)	(0.263)	(0.263)	(0.282)
Observations	6,514	6,514	6,514	6,299	6,299	6,299
R-squared	0.774	0.122	0.145	0.691	0.103	0.134
Number of code	281	281	281	421	421	421

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4.3_ Panel B fixed effect for subsamples- Regional Cities

	East cities			North-east cities		
	m1	m2	m3	m1	m2	m3
VARIABLES	lnpop	lngrowth	lngrowth	lnpop	lngrowth	lngrowth
L.lnpop	0.875***	-0.125***	-0.518***	0.776***	-0.224***	0.322
	(0.0161)	(0.0161)	(0.163)	(0.0790)	(0.0790)	(0.238)
L_1lnpop2			0.0302**			-0.0520**
			(0.0122)			(0.0240)
Constant	0.849***	0.849***	2.108***	1.437***	1.437***	0.0998
	(0.107)	(0.107)	(0.540)	(0.504)	(0.504)	(0.565)
Observations	4,932	4,932	4,932	1,694	1,694	1,694
R-squared	0.813	0.082	0.091	0.642	0.130	0.168
Number of code	284	284	284	88	88	88

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

	Midland cities			West cities		
	m1	m2	m3	m1	m2	m3
VARIABLES	lnpop	lngrowth	lngrowth	lnpop	lngrowth	lngrowth
L.lnpop	0.825***	-0.175***	-0.462**	0.773***	-0.227***	-0.404
	(0.0352)	(0.0352)	(0.233)	(0.0451)	(0.0451)	(0.343)
L_1lnpop2			0.0230			0.0146
			(0.0183)			(0.0264)
Constant	1.137***	1.137***	2.023***	1.416***	1.416***	1.934*
	(0.224)	(0.224)	(0.748)	(0.276)	(0.276)	(1.100)
Observations	2,986	2,986	2,986	3,201	3,201	3,201
R-squared	0.747	0.118	0.121	0.722	0.184	0.188
Number of code	152	152	152	178	178	178

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4.3_ Panel C fixed effect for subsamples- Historical Cities

	29 cities (1879-2009)			82 cities (1936-2009)		
	m1	m2	m3	m1	m2	m3
VARIABLES	lnpop	lngrowth	lngrowth	lnpop	lngrowth	lngrowth
L.lnpop	0.905***	-0.0953***	-0.265	0.909***	-0.0905***	-0.304
	(0.0316)	(0.0316)	(0.282)	(0.0184)	(0.0184)	(0.203)
L_1lnpop2			0.0117			0.0151
			(0.0191)			(0.0143)
Constant	0.742***	0.742***	1.349	0.672***	0.672***	1.412*
	(0.234)	(0.234)	(1.035)	(0.130)	(0.130)	(0.718)
Observations	745	745	745	2,107	2,107	2,107
R-squared	0.837	0.054	0.056	0.848	0.053	0.056
Number of code	29	29	29	82	82	82

	125 cities (1958-2009)			294 cities (1984-2009)		
	m1	m2	m3	m1	m2	m3
VARIABLES	lnpop	lngrowth	lngrowth	lnpop	lngrowth	lngrowth
L.lnpop	0.915***	-0.0848***	-0.207	0.876***	-0.124***	-0.264*
	(0.0148)	(0.0148)	(0.170)	(0.0152)	(0.0152)	(0.156)
L_1lnpop2			0.00890			0.0110
			(0.0123)			(0.0117)
Constant	0.610***	0.610***	1.023*	0.824***	0.824***	1.265**
	(0.102)	(0.102)	(0.586)	(0.0977)	(0.0977)	(0.513)
Observations	3,156	3,156	3,156	6,855	6,855	6,855
R-squared	0.861	0.051	0.052	0.813	0.080	0.082
Number of code	124	124	124	294	294	294

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

The sample uses 29 historical cities existing from 1879, which experienced from the end of Qing Dynasty. The sample uses 82 historical cities existing from 1936, which experienced from the Republic of China- ruled by nationalist party-. The sample uses 294 cities existing from 1984, which experienced from the start of 'Economic Reform'. The sample uses 125 historical cities existing from 1958, which experienced from the PRC (People's Republic of China)- ruled by communist party-. The sample uses 294 cities existing from 1984, which experienced from the start of 'Economic Reform'.

4.4.3 Panel Unit Root Test for Gibrat's law

Then we notice that the system GMM test is equivalent to a panel unit root test for city size, if there is unit root then the process behind city growth is stochastic and Gibrat's law holds. From Table 4.4 below we find that the null hypothesis which supposes that all panels contain a unit root is highly rejected, that means that at least one panel is stationary which indicates Gibrat's law does not hold for all the cities. This result is consistent with the system GMM and the correlation analysis we analysed above. Panel Unit root test for subsamples like prefecture-level cities and county-level cities; East cities, Midland cities, West cities and North-East cities are reported in Appendix Table A4.3, almost all highly reject the null hypothesis that all panels contain unit root, i.e. at least for some cities Gibrat's law does not hold.

Table 4.4 Panel unit root test- Group 1 Whole sample

Panel Unitroot Test		(1) ADF Test Without trend	(2) ADF Test With trend	(3) ADF (2) Test Without trend	(4) ADF (2) Test With trend	(5) PP Without trend	(6) PP With trend	(7) PP (2) Without trend	(8) PP (2) With trend
Inverse chi-squared	P	2916.9041***	2953.0736***	1969.8133***	2266.3196***	5318.9262***	5543.6169***	5380.1776***	5737.4551***
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Inverse normal	Z	-0.1659	-2.4916***	9.3250***	4.8567	-14.8649***	-18.5994***	-15.6699***	-20.3082***
		(0.4341)	(0.0064)	(0.0000)	(1.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Inverse logit t	L*	-7.9533***	-11.9727***	5.5071***	-2.3689***	-34.4083***	-41.6886***	--35.3884 ***	-44.3755***
		(0.0000)	(0.0000)	(0.0000)	(0.0090)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Modified inv. chi-squared	P_m	28.8421***	29.5271***	11.4123***	17.0726***	74.1515***	78.4008***	75.3099***	82.0666 ***
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Number of panels		707	707	707	707	707	707	282	282
Avg. number of periods		21.21	21.21	21.21	21.21	21.21	21.21	27.70	27.70

Ho: All panels contain unit roots. Ha: At least one panel is stationary. P-values are reported in parentheses. The null hypothesis is that the variable contains aunit root, and the alternative is that the variable was generated by a stationary process. ('trend' specifies that a trend term be included in the associated regression and that the process under the null hypothesis is a random walk, perhaps with drift. This option may not be used with the no constant or drift option.)

ADF refers to Augmented Dickey-Fuller test and PP test refers to Phillips-Perron test, 1988. Other panel unit roots tests are not available due to our unbalanced data.

Testing for Gibrat's law is equivalent to testing for unit roots in the evolution of city sizes. If there is unit roots, thus the data is not stationary which indicates Gibrat's law holds.

4.4.4 Test for Structural Break

There are two recent influential policies in China that could significantly affect the growth of cities- 'China Economic Reform' and 'One Child per Family Policy' which launched almost at the same time in the end of the 1970s and were effective from the early 1980s. With the 'Economic Reform' beginning in 1979, China opened up market economy, loosened the government control and the restriction on intercity migration; abandoned the policy of deliberate city development in the interior; almost coincidentally it also introduced the 'One Child Family Policy', especially in the cities where the policy was most effectively monitored. Therefore a structural break in the process of city sizes is to be expected around the late 1970s. Then we add year dummies to investigate the structural break. Due to the availability of our data and Zipf's exponent evolution analysis (most of the turning point is around 1984), we choose the year 1984 as a year dummy, since for the first few years these two policies need time to take effect.

Results are reported in Table 4.5 below, although the dummy variable d1984 is significant at 1% level, the estimated coefficients of other variables are the same with the results with no dummy added sample (Table 4.2). This is consistent with above findings about Zipf's law and Gibrat's law that 'Economic Reform' or 'One Child Policy' does not truly affect the validity of these two regularities.

Table 4.5: Structural Break

Add dummy 1984 to whole sample, without robust standard errors (use command xtdpdsys, twostep)

	(1)	(2)	(3)	(4)	(5)
	m1	m2	m3	m4	m5
VARIABLES	lnpop	lngrowth	lngrowth	lngrate	lngrate
L.lnpop	0.647***	-0.618***	-1.938***	-1.111***	1.087***
	(2.76e-05)	(0.000213)	(0.00163)	(2.62e-07)	(2.39e-07)
L_1lnpop2			0.101***		-0.172***
			(0.000134)		(3.75e-08)
L.lngrowth		-0.0606***	-0.0597***		
		(3.90e-05)	(4.45e-05)		
L.lngrate				-4.11e-07***	-4.88e-07***
				(0)	(0)
d1984	1.932***				
	(0.564)				
Constant	0.355	3.984***	8.231***	7.160***	0
	(0.564)	(0.00207)	(0.00485)	(1.63e-06)	(0)
Observations	12,850	11,680	11,680	11,680	11,680
Number of cities	702	699	699	699	699
Instruments	827	367	368	367	368
Test d1984=0	0.0006				
Test β_1=1	0.0000				
AR(1)	0.0000	0.0000	0.0000	0.0000	
AR(2)	0.6776	0.3973	0.8733	0.7847	
Sargan	1.0000	0.0012	0.0012	0.0002	

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
P-values are reported in the tests of d1984=0, β_1=1, AR tests and Sargan tests.
Note that d1984 dropped because of collinearity in model 2 to model 5, and the constant omitted in model 5 thus cannot calculate AR tests and Sargan test with dropped variables.

4.5 CONCLUSION

With respect to Gibrat's law, it is found to be approximately held for the whole sample during 1980s to 2000s which can be considered as the explanation for Zipf's law which appear to hold in the end of the 1980s and early 1990s. This is consistent with Gabaix's (1999) explanation for Zipf's law that under Gibrat's-mode of urban growth, Zipf's law will show up at the steady state.

In addition, for the first time, we divide the whole sample several times as prefecture-level cities and county level cities; East, Midland, West and North-East cities; 4 historical city groups. (1) Gibrat's law is more supported in prefecture-level cities and rejected in county-level cities, which also consistent with previous Zipf's results that Zipf's law showed in prefecture-level cities' distribution and never showed in county-level cities. County level cities are distributed less evenly as Zipf's law predicts, i.e. the disparity between the size of large and small cities is big. (2) Gibrat's law is found in East region cities, which is also consistent with Zipf's results for East regions which shows that the Pareto exponent equals to -1.029, i.e. Zipf's law holds. In contrast, the other three regions do not follow Gibrat's-mode of growth thus do not show Zipf's law. (3) from the historical cities sample, we can see that the youngest city group d- 294 cities (1984-2009) is the closest to Gibrat's law, while other older historical cities are slowly moving towards Gibrat's mode of urban growth. Before the early 1980s, large cities seem to grow more slowly than smaller ones. Since the 'Economic Reform' (pursuing 'development first equality second') favours large cities in the first stage, we observe large cities beginning to grow faster than smaller ones and then the growth rate becomes independent of initial size gradually. This indicates that Chinese cities are growing parallel to each other.

CHAPTER 5
CHINESE CITY GROWTH: TESTING FOR SEQUENTIAL CITY GROWTH

5.1 INTRODUCTION

The mode of city formation and city growth has long been the main interest of urban economics. Chapter 2 and 3 explore the city size distribution in China by testing for Zipf's law and Gibrat's law and find that the evolution of Chinese city size distribution has passed the distribution that Zipf's law described and grew toward Gibrat's law during 1879 to 2009 (Zipf's law holds for a few years in late 1980s and early 1990s, but none of them holds strictly). These two chapters study the spatial agglomeration in a static context, i.e. Zipf's law and Gibrat's law describe the urban hierarchy at a given point of time. However, in this chapter, we will extend our Chinese urban analysis to a dynamic model- whether Chinese city growth follows the Sequential city growth theory, Cuberes (2011). This theory claims that within a country the largest city will grow the fastest initially, and then as time passes the second largest city will become the 'fastest grower', and then at some point eventually the third-largest city starts growing the fastest and so on; the rank of the 'fastest grower' is moving from the large cites (rank1) to the small cities (rank n) along urban hierarchy over time.

Furthermore, despite the large amount of literature quantifying the city size effect on growth, including the testing of Gibrat's law (some studies support Gibrat' such as Ioannides and Overman, 2003; Eeckhout, 2004 for US; Giesen and Sudekum, 2010 for Germany; while others reject such as Black and Henderson, 2003 for US; Bosker *et al.*, 2008 for West Germany), and the sequential city growth theory (Cuberes, 2009) plus its empirical findings (Cuberes, 2011), there is few evidence of the impact of a city's age on its growth. Therefore, we also test the age-dependent patterns of urban growth, Sanchez-Vidal *et al*. (2014). The original sequential city growth theory addresses that the largest-size city will grow the fastest initially, and then the second largest-size city, and so on. However, age sequential city growth consider the age of a city rather than the size, and argues that the young city will grow the fastest initially and then when they are mature their growth rates tends to slow down or even decline.

We use the same data as previous chapters- the population data for Chinese cities from 1879 to 2010 (we update 2010 population data in this chapter, the number of cities varies over year) to test for sequential city growth and also the age-dependent patterns of urban growth, i.e. the age sequential city growth. For the size sequential, we find that firstly, the distribution of city's growth rate is skewed to the right for the whole period in China, i.e. from 1890 to 2010 for each year only a few cities grow much faster than the rest. Secondly, the growth pattern tends to follow sequential city growth as shown in Cuberes (2009, 2011) if we use the same OLS method. We calculate the average rank of the first quartile of fastest growing cities, and then we analyse whether this average rank increases over time. If this is true, then it indicates that the largest city (with the smallest rank) grow the fastest initially, and then as time passes, the second largest city is the fastest grower and so on, and then the fast-grower can be found in the third largest city in the next period, until to the small cities (with large ranks).

However, the driving force of the increasing of the average rank of the fast-growers may come from the increasing number of cities each year. To control for this factor, as mentioned in Cuberes (2011) that we add the number of cities (N_t) in the regression equation, we also test for some subsamples. Firstly, for different time spans sample we also test for 1936-2010, 1953-2010 and 1984 to 2010. Secondly, we test for a sample of large cities comprising the top 300 large cities for each year from 1985 to 2010, to investigate the growth pattern in large cities. Lastly, we test for the fixed 197 cities which exist every year during 1985 to 2010. For the last two samples the growth rate for the number of cities each year is 0%, because the number of cities for each year is fixed.

To test for the age sequential city growth pattern, we create several age dummies according to Sanchez-Vidal *et al*. (2014) and we find "inverse" age sequential city growth in Chinese cities. This indicates that unlike Sanchez-Vidal *et al*. (2014)'s finding that in the first few decades when cities emerge (young cities), they tend to grow the fastest and then when they become more mature the growth rate slows down, we find the inverse growth pattern according to city's age. We find that in China older cities tend to grow the fastest initially, and then as time passes the growth rate slows down. We also test for the prefecture-level cities and four different economic regions. Results are basically the same as the whole sample, only for the north-eastern region, in general cities grow faster than other regions and the more mature the faster the growth.

Our study in this chapter contributes to both the urban designers and academics. Firstly, the novel empirical findings in China are helpful to form effective policies, such as infrastructure investment, especially in early rapid urbanisation stage. A large body of literature on urban economics emphasise the essential role of infrastructure investment in economic performance. We use Chinese city-level data to present how urban development evolves over time, thus contribute to the design of strategies on city infrastructure investment, welfare and health system, etc. Secondly, this chapter contributes to the urban economics literature in terms of understanding the city growth and especially the impact of urbanisation process on its urban structure in China. Furthermore, the empirical findings in China would probably be valuable new evidence in terms of extending or modifying urban growth theories.

This chapter is based on the theoretical and empirical analysis of sequential city growth of Cuberes (2009, 2011), however, there are several important differences. Firstly, in Cuberes (2011), he use a panel data of 54 cities globally from late 19th century to early 21st century (the time period is on a decade frequency) including Chinese cities from 1890-1994. While, in this chapter we analyse Chinese cities only, and extend time period over one decade, to 2010 (decade frequency before 1984 and yearly frequency after 1985, inclusive). Secondly, Cuberes's paper only did 0.6 cut-off sample[39], the average growth rate of the number of cities over year in his sample across 54 countries is about 12%, and in China it is about 14% in his chapter.

5.2 LITERATURE REVIEW

[39] Cuberes (2011) select cities that have a relative population above 0.6, this threshold comes from dividing the US median city size in 1790 (5077) by its average (8402). This method follows Henderson and Wang (2007) who argue that this sample selection method can allow one to analyse a portion of the city size distribution that is comparable across countries and over time.

5.2.1 Size Sequential

The formation and development of cities was first studied in a static context, for instance why cities exist in different types at a given point (Henderson, 1974). And then Fujita (1976) initially study the spatial agglomeration in a dynamic way. Ana (1978, 1992), Kanemoto (1980), Henderson and Ioannides (1981), Miyao (1981), Fujita (1982), Krugman (1992), Ioannides (1994), and Palivos and Wang (1996) [40]are also important contributions to study the dynamic spatial agglomeration.

The drawbacks of the previous literature on urban growth is that they assume free mobility of all factors of production, thus they predict firstly, large and rapid swings in city population that reach a critical level; secondly new cities population jumps instantly to some arbitrarily size. These results are obvious contrary to the existing data which has smooth fluctuations, Henderson and Venables (2009).

Then several studies attempt to improve the model and generate smooth changes of cities' population. One attempt of solving this problem is that of Brezis and Krugman (1997) who propose a model that cities grow sequentially forced by technological innovations. The introduction of new technology stimulates a rapid increase of population in the original city, but after a period of time, its population declines substantially while the new city's population rises by increasing relative productivity through learning. Although this seems a plausible theoretical mechanism, there are still doubts about whether the main force driving the city population dynamics is the launch of new technology. Another

[40] Duranton and Puga (2004) and Rossi-Hansberg and Wright (2007) reviewed these literature thoroughly.

attempt is to premise that one of the production factors is immobile. A limited number of researches have studied this and addressed the irreversible investment in capital good. Fujita (1978) assumes capital immobility and proposes an optimal dynamic equilibrium, but with no decentralized version.

Currently only three papers address the sequential city growth model explicitly, Henderson and Venables (2009) and Cuberes (2009, 2011). Henderson and Venables (2009) analyse city formation within a country where cities' population grows steadily over time. In contrast to most of the previous literature, they assume a) immobile housing and urban structure b) agents are forward looking, with fixed assets in the city, a sequential pattern of city formation has been generated. It predicts a smooth change in urban population with swings in house rent, which is unlike previous literature presenting rapid fluctuations in urban population. Henderson and Venables also address the role of institutions in intervention by local government.

Cuberes (2009) complements Henderson and Venables (2009) in important ways. It highlights the role of physical capital which is an input in each firm's production function rather than residential capital in Henderson and Venables (2009). Specifically, it proposes a dynamic growth model to optimal city size that generates a simple mechanism where cities grow sequentially, which rationalizes the sequential city growth pattern. According to the model, at any point in time, one city grows much faster than the rest, and the rank of the fastest growing cities increases as time goes by. The model also predicts that sequential city growth is faster in periods of rapid urban population growth. In his model the changes of capital and population in two cities over time (assume two cities in the country, and then assume three cities) are in Figure 5.1 as follow.

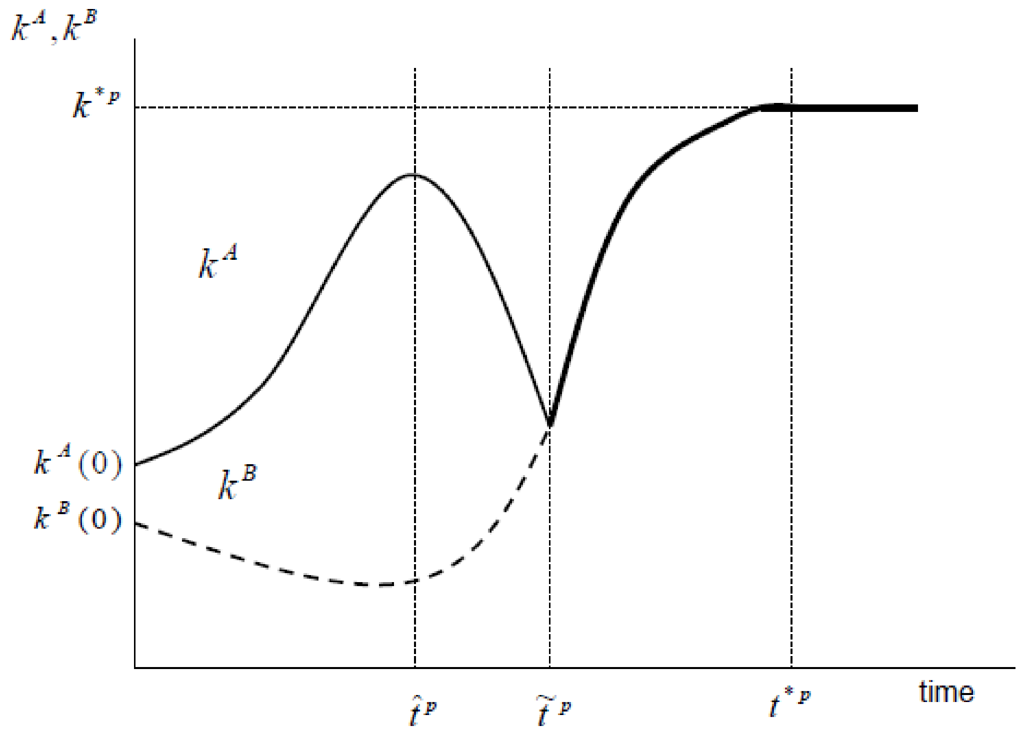


Figure 5.1-a Capital change in two cities, Cuberes (2009)

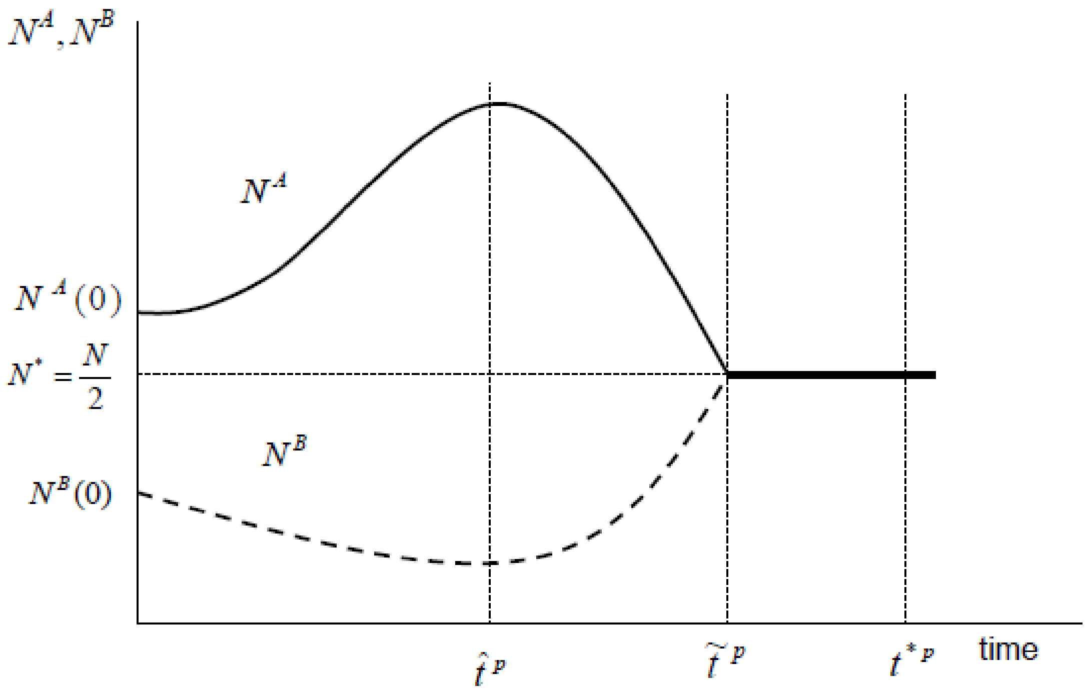


Figure 5.1-b Population change in two cities, Cuberes (2009)

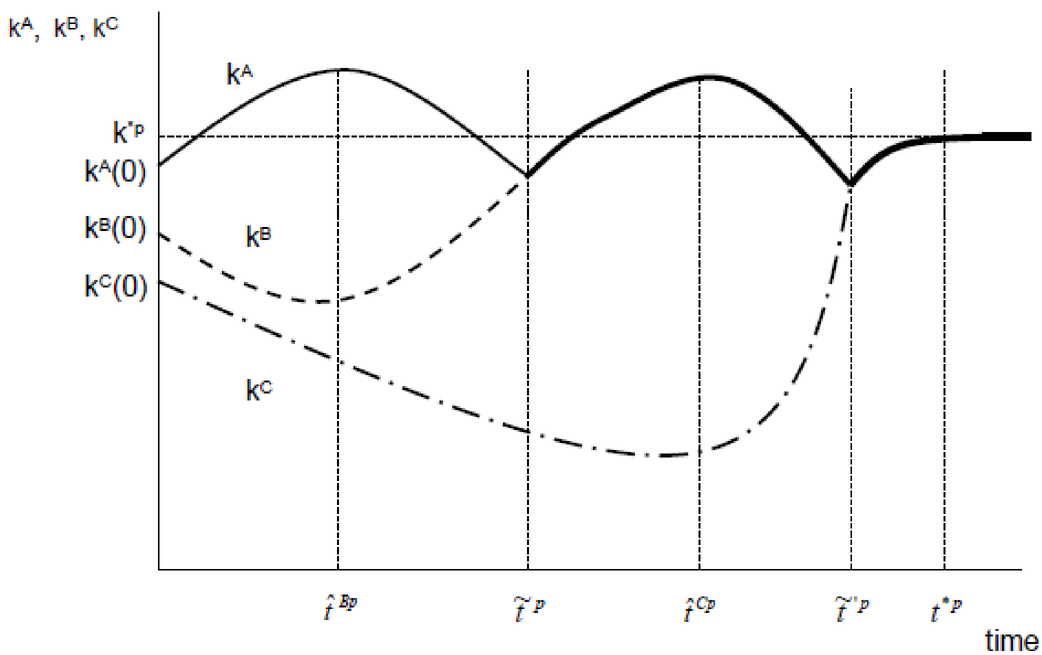


Figure 5.1-c Capital change in three cities, Cuberes (2009)

We summarise Cuberes (2009)'s model for sequential city growth as below. Cuberes (2009)'s model assumes that there are two cities in a country modelled as Cobb-Douglas production functions. Labour and capital are the only two factors to produce a homogenous product for each city. Output Y^{ij} of firm i who located in city j can be expressed as

$$Y^{ij} = \left(N^{ij}\right)^{\alpha}\left(K^{ij}\right)^{1-\alpha}(K^{j})^{\varphi} \qquad (1)$$

Where N^{ij} denotes the labour inputs for firm i in city j, similarly K^{ij} refers to the capital for firm i in city j. Thus, $K^{j} = \sum_{i=1}^{I} K^{ij}$ represents the total capital stock in city j. With $1 > \alpha > 0$ and $\varphi > 0$, firms located within the city have positive external effect. Furthermore, the profit for a firm can be expressed as following, when normalizing the price of the production to one, given that firms would pay a fraction of congestion cost $g\left(K^{j}\right)$ produced by the total capital stock in the city.

$$\pi^{ij} = \left(N^{ij}\right)^{\alpha}\left(K^{ij}\right)^{1-\alpha}(K^{j})^{\varphi} - \left(r^{j} + \delta\right)g\left(K^{ij}\right) - \frac{1}{I}g\left(K^{j}\right) - \omega N^{ij} \quad (2)$$

where $g(.)$ is an increasing and convex function; $\delta \in (0,1)$ denotes the capital depreciation rate, and r^{j} and ω represent the price of capital and the wage rate, respectively. If assume free labour mobility, then we can conclude that the two city's population ratio has the following relationship:

$$\pi^{iA} = \pi^{iB} \tag{3}$$

i.e.

$$\frac{N^A}{N^B} = (\frac{K^A}{K^B})^{\frac{1+\varphi-\alpha}{1-\alpha}} \tag{4}$$

With the firm's first-order conditions, one can also have

$$r^j = f_j - \delta - g'(K^j) \tag{5}$$

Given that the gross marginal product of capital in city j is $f_j \equiv (1-\alpha)\left(N^{ij}\right)^{\alpha}\left(K^j\right)^{\varphi-\alpha}$, and households invest in capital and supply labour inelastically. They solve the following optimisation problem (Cuberes, 2009):

$$max \int_0^{\infty} e^{-\rho t} \ln(c)\, dt \tag{6}$$

$$\sum_{j=A,B} i^j + c = \omega + \sum_{j=A,B} r^j z^j \tag{7}$$

$i^j \geq 0, \forall j = A, B$

z_0^j given, $\forall j = A, B$

where z^j is the assets value invested in city j, c is per-capita consumption, and $\rho \in (0,1)$ is the household's discount rate. The irreversible capital constraints is a very important assumption in the model, each household faces the constraint: $i^j \geq 0, \forall j = A, B$. This restricts the relocate or destroy of the capital to other cities once invested in one city. And the initial assets stock in city j has been given- z_0^j. Next, the model assumes the capital stock in city A is a bit larger than city B at initial time, thus city A's congestion costs can be considered small compared to the productivity gains from this large size of capital stock. Under these assumptions, for these two cities A and B, the growth pattern of each city evolves in Figure 5.1-b.

In Figure 5.1-a, we can observe that initially as capital stock is relatively larger in city A than city B, thus the population size is larger in city A. According to equation (4), in the model population moves following the capital. Therefore, during the initial period to period $\hat{t}^p$, population would move from city B to city A (from small city to large city) as all of the new investment moving into city A. And then, at time $\hat{t}^p$ capital stock would be productive equally between city A and B as the increasing congestion cost in city A during initial period to period $\hat{t}^p$. And the investment inflows into city B until there shows the same capital stock level in city A and B (at period $\tilde{t}^p$) . After this growing period, the two cities tend to be identical and the population is distributed equally between city A and B, until the steady state is reached in the economy (at periodt^{*p}).

The figures above from Cuberes (2009, 2011) clearly illustrate the sequential city growth, which at current period the largest city grows the fastest initially, and then as time passes when the largest city grows to its critical value, the initially second-large city grows the fastest; as time goes by, after the second large city reaches its critical level the third large city will become the ‘fastest grower’.

Cuberes (2011) tests three new empirical facts using the model above proposed in Cuberes (2009): a) the distribution of the growth rate of city population is right skewed, which indicates only a few cities will have high growth rate of population in each point in time during the transition to the steady state; b) with in a country or an economy, the average rank of fastest growing cities in each point in time increases over time; c) this sequential growth process is more significant when urban population grows rapidly. It uses a long time period historical data for 54 countries' administratively defined cities and metropolitan areas and finds that for most of the countries cities tend to grow sequentially.

5.2.2 Age Sequential

Although there is large body of literature emphasising the impact of city size (population size) on city growth, there is few evidence to show whether there is impact of city's age on its growth. Recently Sánchez-Vidal *et al.* (2014) provide some new empirical evidence of sequential city growth with respect to city age, using US incorporated place-level data from 1900 to 2000. They argue that the young small cities grow at a faster rate initially, but as decades pass their growth rates tend to slow down and even decline. This inspires us to test for whether age matters in urban growth patterns in China, using the same data as size sequential city growth in above section.

The inclusion of new (born) cities is also analysed in previous works by Dobkins and Ioannides (2000) and Henderson and Wang (2007). They define their new (born) cities in their datasets when a city's population reaches a critical threshold. Nonetheless, Giesen and Sudekum (2012) include all of the new cities without any threshold restrictions. They propose a theoretical model and find that the distribution of city size within a country is significantly associated with the city's age distribution. They argue that initially young cities grow the fastest; however, in the long run all of the cities grow at the same rate (Gibrat's law). Desmet and Rappaport (2013) find that smaller counties tend to converge in earlier periods while larger ones tend to diverge, however, as city' age composition changes over time within a country, both the convergence and divergence pattern faded out and Gibrat's law gradually emerges.

In this chapter, we attempt to test for whether there is a sequential city growth pattern in Chinese cities according to their city age. Recall that the sequential city growth refers to within a country, a few cities initially grow much faster than the rest (the growth distribution skewed to the right), and however, as time passes their growth slows down and other cities begin to grow faster than the rest ones, and so on (Cuberes, 2009; Henderson and Venables, 2009).

The only empirical study to this sequential city growth theory is Cuberes (2011) until now, who valid the sequential city growth theory by panel data consisting of cities from 54 countries and alternatively metropolitan areas from 115 countries (time period varies in different countries). Cuberes (2011) shows that cities within a country do show a sequential growth pattern, i.e. large cities initially grow faster and then the second large cities grow faster in its turn etc.. His study shows the sequential city growth driven by the size (population size) of cities, while this section we use the city's age instead of city size we test for the age-dependent sequential growth model.

5.3 DATA AND METHODOLOGY

5.3.1 Data

5.3.1.1 The data

This paper uses Chinese administrative city level data from 1879 to 2010 which is combined from two data sources. The historical data for annual city population before 1984 (1879 to 1983) comes from the most comprehensive dataset of world urban populations, by Jan Lahmeyer *(http://www.populstat.info/)*. (It is the same data as Cuberes 2011 used for China.) This data is consistent with city proper -*'Shiqu'*- data in *'China Urban Statistical Yearbook'* as well. However, the historical population data in some years is not complete due to certain external conditions (wars) and census technologies during that period.

Chinese city level annual population data from 1984 to 2009 are provided by the '*China Urban Statistical Yearbooks' (National Bureau of Statistics of China, 1985-2010).* We use the city-proper data[41] which is the population of inner urban area, not including the subordinate towns. Data in the yearbook series goes back to 1984 and the city definitions

[41] Details about city-proper definition in China can be found in Zipf's Chapter.

obey the 1984 criteria[42], adjusted to year 2000 statistical changes. Hong Kong, Macao and Taiwan are excluded.

Table 5.1 below shows the description statistics of the whole sample with zero-cut off. Figure 5.1- Panel A shows that coefficients of skewness of the population growth rate for each year, which are mostly concentrated in positive values. This shows that city growth rate is skewed to the right, which indicates that 'fast growers' are minority, i.e. only a few cities have relatively fast growth rate. Figure 5.1- Panel B shows the frequency of population growth rate for year 2010, which also prove that only a few cities experience fastest growth. Figure 5.2- Panel A and Panel B shows the time trend of the average rank of the 'fast-growers' in absolute value and scaled value respectively. From the first graph, the average rank of the 'fast-growers' is increasing over time from 1890 to the mid-1990s, which is consistent with the prediction of sequential city growth, but after the mid-1990s the average rank seems stable and decrease a bit until 2010. In the second graph, there seems a reversion of the average rank to the 0.5, which is not consistent with the sequential city growth. These will be further tested in the following sections.

[42] From 1983 Chinese urban system launched the policy of 'city governing the surrounding towns', the central government assigned some towns to the adjacent city in order to take the advantage of the city to help the development of the towns.

Table 5.1: Data description- Whole sample- zero cut-off

Year	no. of cities	no. of cities growth rate	no. of matching cities	Skewness of growth rate	rank25	Growth rate rank25	Min rank25	Max rank25	Urban pop	Urban pop growth rate	Urbanisation rate	National total pop
1879	31		16						11.4004		3.11%	366.988
1890	19	-38.71%	16	0.2279	7		1	13	6.3366	-44.42%	1.67%	380
1900	25	31.58%	17	4.0293	18	157.14%	9	22	8.1718	28.96%	2.04%	400
1911	38	52.00%	23	1.1755	17.17	-4.61%	5	30	9.7625	19.47%	2.28%	427.662
1918	30	-21.05%	17	0.5463	8.8	-48.75%	1	21	8.0167	-17.88%	1.74%	461.766
1926	31	3.33%	27	2.7808	19.57	122.39%	13	30	11.6375	45.17%	2.41%	482.128
1936	82	164.52%	24	0.4196	16.83	-14.00%	1	33	23.7842	104.38%	4.68%	507.864
1948	88	7.32%	82	3.3855	37.05	120.14%	8	77	31.373	31.91%	5.73%	547.804
1953	153	73.86%	88	6.7856	45.24	22.11%	4	89	49.7993	58.73%	8.47%	587.96
1958	127	-16.99%	92	3.1034	32.83	-27.43%	2	81	65.1313	30.79%	9.96%	654.159
1983	195	53.54%	116	2.0261	69.41	111.42%	10	181	222.74	241.99%	21.62%	1030.08
1984	207	6.15%	187	3.6896	114.98	65.65%	3	204	240.17	7.83%	23.01%	1043.57
1985	324	56.52%	294	10.0169	213.22	85.44%	15	324	250.94	4.48%	23.71%	1058.51
1986	321	-0.93%	301	10.1224	205.09	-3.81%	5	321	263.66	5.07%	24.52%	1075.07
1987	382	19.00%	318	6.9744	243.91	18.93%	2	382	276.74	4.96%	25.32%	1093
1988	434	13.61%	378	12.5915	280.38	14.95%	16	431	286.61	3.57%	25.81%	1110.26
1989	449	3.46%	430	12.2407	281.61	0.44%	1	449	295.4	3.07%	26.21%	1127.04
1990	467	4.01%	446	9.9770	259.35	-7.90%	14	466	301.95	2.22%	26.41%	1143.33
1991	478	2.36%	464	-9.8751	306.91	18.34%	18	475	312.03	3.34%	26.94%	1158.23
1992	517	8.16%	475	17.4947	334.81	9.09%	23	516	321.75	3.12%	27.46%	1171.71
1993	570	10.25%	465	11.5709	368.45	10.05%	1	569	331.73	3.10%	27.99%	1185.17
1994	620	8.77%	563	12.3405	374.04	1.52%	6	619	341.69	3.00%	28.51%	1198.5
1995	638	2.90%	614	6.5742	390.49	4.40%	4	637	351.74	2.94%	29.04%	1211.21

1996	664	4.08%	635	23.9774	395.58	1.30%	8	663	373.04	6.06%	30.48%	1223.89
1999	664	0.00%	655	7.5846	355.76	-10.07%	1	664	394.49	5.75%	31.91%	1236.26
2000	653	-1.66%	643	7.4429	343.19	-3.53%	3	650	416.08	5.47%	33.35%	1247.61
2001	664	1.68%	650	14.0443	333.99	-2.68%	1	664	480.64	15.52%	37.66%	1276.27
2002	656	-1.20%	647	15.1689	343.99	2.99%	2	656	502.12	4.47%	39.09%	1284.53
2003	655	-0.15%	648	10.3590	315.28	-8.35%	4	655	523.76	4.31%	40.53%	1292.27
2004	654	-0.15%	646	15.3146	314.44	-0.27%	6	654	542.83	3.64%	41.76%	1299.88
2005	658	0.61%	650	12.8599	322.74	2.64%	2	657	562.12	3.55%	42.99%	1307.56
2006	660	0.30%	656	2.4054	316.45	-1.95%	1	660	582.88	3.69%	44.34%	1314.48
2007	654	-0.91%	653	0.9553	336.4	6.30%	5	653	606.33	4.02%	45.89%	1321.29
2008	655	0.15%	654	9.0581	342.01	1.67%	3	653	624.03	2.92%	46.99%	1328.02
2009	654	-0.15%	654	11.6994	322.65	-5.66%	3	654	645.12	3.38%	48.34%	1334.5
2010	656	0.31%	651	1.5231	319.25	-1.05%	8	655	669.78	3.82%	49.95%	1340.91
Avg.	410.36	12.76%	385.97	7.4454	228.768	18.73%	5.971	414.514	304.04954	17.50%	25.33%	1006.374

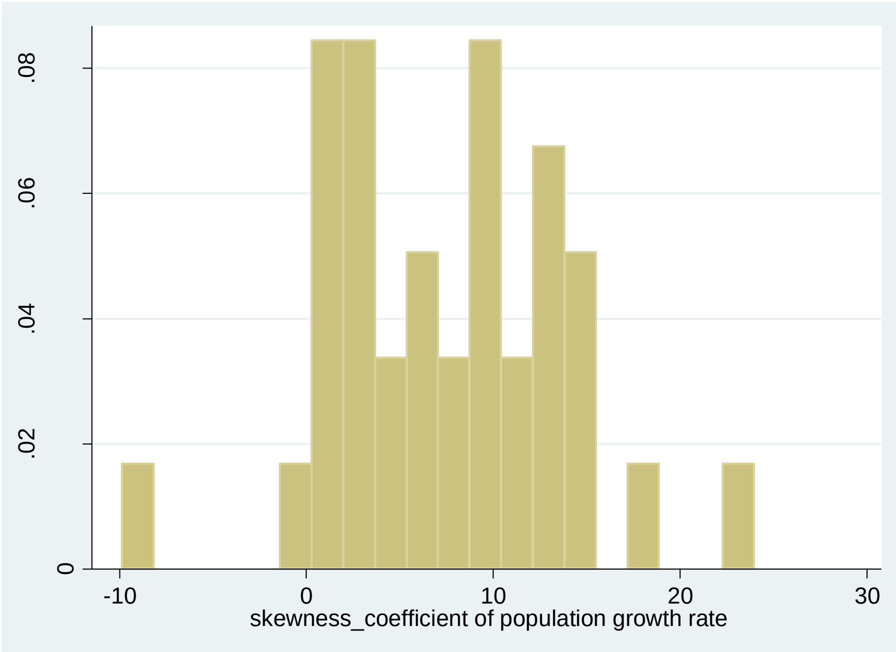


Figure 5.2- Panel A

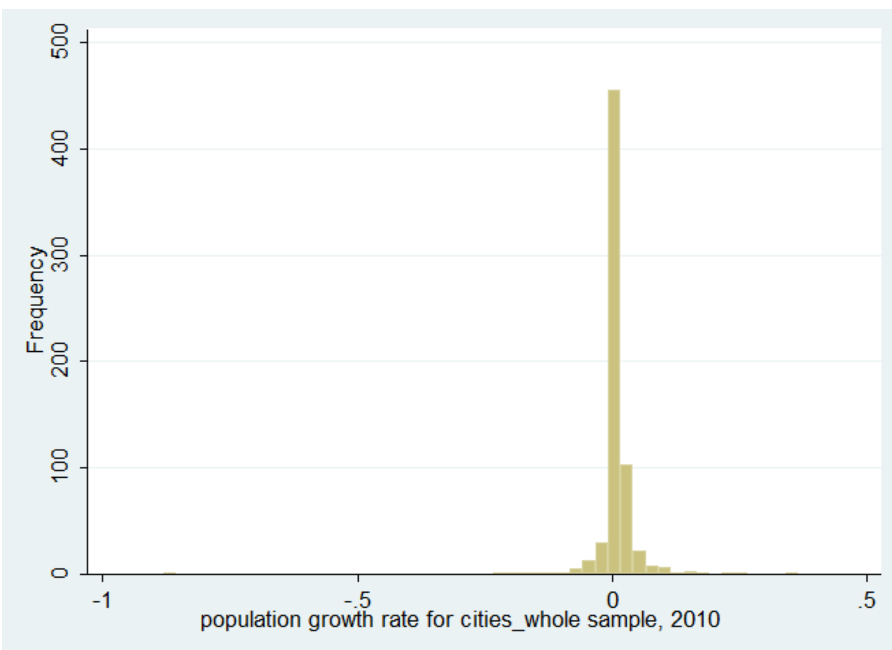


Figure 5.2- Panel B

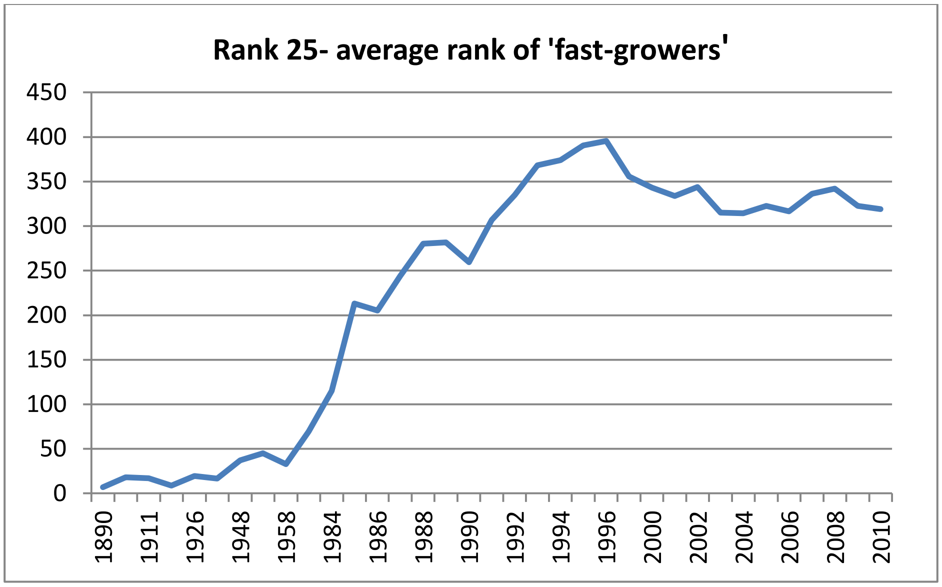


Figure 5.3- Panel A

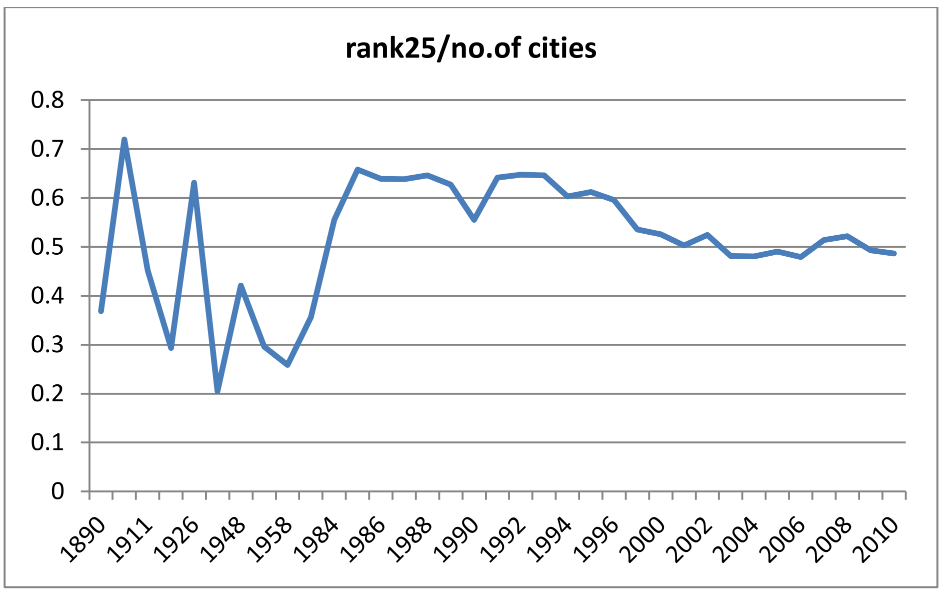


Figure 5.3- Panel B

5.3.1.2 Sample Selection of Cities

In this paper, as our historical data has a long time span and due to the heterogeneity of city definitions and statistical method of population in different periods, we follow the same methodology used in Cuberes (2011). He set up a cut-off when selecting cities, i.e.

if cities are ordered by their population; one selects the first s cities, thus the $s + 1$ city would be below a relative cut-off. This threshold [43] in Cuberes (2011) paper comes from dividing the US median city size in 1790 (5077) by its average (8402) which is 0.6, i.e. it selects the cities that have a relative population over 0.6. In our sample, we use the threshold as 0.7 which comes from dividing Chinese median city population in 1879 (276,600) by its average (393,117).

For the data testing for the age-sequential growth pattern, we use the city population for each decade from the same data source. From the 1890s to the 1980s since the data is not available for every year we use the average population for each city during one decade. For the 1990s and 2010s we use the population data for 1990 and 2010, respectively.

5.3.2 Methodology

5.3.2.1 Testing for Size-Sequential Growth

The procedure of testing for sequential city growth follows Cuberes (2011). Firstly, for each year, we rank the city size by population decreasingly: with the largest city having rank 1 and the second rank 2 etc. Next, we calculate the growth rate of each city between current year and the previous year from 1984 to 2010, while from 1879 to 1983 we

[43] This sample section method actually comes from Henderson and Wang (2007), the cut-off is defined as the ratio of the minimum (100,000) over mean (495,101) city size (population) in their sample of countries in 1960. Henderson and Wang suggest that the advantage of this sample selection method is that when researchers attempt to compare results across countries and over time, this analyse approach can address a portion of the city size distribution and make it possible to compare.

calculate the decade growth rate of each city due to the data availability during economically and politically unstable periods. The same as Cuberes (2011), we set the third quartile of city growth rates for each year as the threshold[44] to find the fastest growing cities for each year. Thirdly, we choose cities whose growth rate is greater or equal to this threshold and consider them as "fast-growers" in that year. Then we calculate the average rank of these "fast-growers" for each year, if sequential city growth theory holds we would expect to find that the average rank might be increasing over time. Besides, within a year, we could also see which part of the urban hierarchy grows the fastest, i.e. do the large cities (lower rank) or the small ones (higher rank) grow the fastest.

More specifically, the procedure could be illustrated as Table 5.2 below which shows the city growth rate for the year 1890. Firstly cities are sorted in descending order by size (population) in 1890, column 4 and 5 show the population in year 1879 and 1890 respectively. The growth rate of each city size between these two years is reported in column 6. Column 7 displays the sample selection method mentioned in Cuberes (2011) who chose 0.6 as the threshold[45] which represent the proportion of each city's population to the mean of the total urban population. Cuberes chooses cities above this cut-off in their sample, similarly in this paper we will choose cities above the threshold of 0.7, as mentioned before.

[44] This threshold has been carried out using different percentiles and the results are very similar.
[45] Cuberes (2011) select cities that have a relative population above 0.6, this threshold comes from dividing the US median city size in 1790 (5077) by its average (8402). This method follows Henderson and Wang (2007) who argue that this sample selection method can allow one to analyse a portion of the city size distribution that is comparable across countries and over time.

Next we will find the cities whose growth rate is strictly greater than the third quartile (75th percentile) of the growth rates according to Cuberes (2011). In the 1890s this third quartile threshold corresponds to the growth rate of 3.23%, so only cities having a growth rate greater than this threshold are classified as 'fast-growers'. Then, we take the average of the rank of these 'fast-growers' for each year or decade ($\boldsymbol{Rank_{25t}}$) and analyse the trend over time, to see whether this average rank increases over time which indicates sequential city growth. Therefore, we estimate the following model as Cuberes (2011) suggests,

$$\log Rank_{25t} = \beta_0 + \beta_1 t + \beta_2 N_t + \beta_3 N_t^2 + \varepsilon_t \quad (8)$$

where $\boldsymbol{Rank_{25t}}$ is the average rank of the top 25% fastest-growing cities; and $\boldsymbol{N_t}$ is the number of cities in each year. In addition, $\boldsymbol{N_t^2}$ is the square term of the number of cities. Variable $\boldsymbol{t}$ measures time in years and $\boldsymbol{\varepsilon_t}$ is a standard error term. We expect that there will be a significant positive sign of coefficient $\boldsymbol{\beta_1}$ if the sequential city growth holds.

Table 5.2: Illustration of calculating procedure

[1]	[2]	[3]	[4]	[5]	[6]	[7]
code	city name	rank in 1890	1879	1890	growth rate of 1890s	pop/avg.in 1890
440100	Guangzhou	1	1,000,000	1,600,000	0.60	4.55
120000	Tianjin	2	930,000	950,000	0.02	2.70
420100	Wuhan	3	800,000	895,000	0.12	2.54
110000	Beijing	4	1,648,800	805,100	-0.51	2.29
350100	Fuzhou	5	630,000	635,000	0.01	1.80
310000	Shanghai	6	276,600	375,000	0.36	1.07
500000	Chongqing	7	250,000	250,000	0.00	0.71
330200	Ningbo	8	115,000	250,000	1.17	0.71
320100	Nanjing	9	450,000	150,000	-0.67	0.43
350200	Xiamen	10	300,000	95,600	-0.68	0.27
330300	Wenzhou	11	500,000	83,000	-0.83	0.24
340200	Wuhu	12	60,000	78,700	0.31	0.22
360400	Jiujiang	13	35,000	53,000	0.51	0.15
420500	Yichang	14		34,000		0.10
440500	Shantou	15	45,000	32,500	-0.28	0.09
450500	Beihai	16	20,000	25,000	0.25	0.07
370600	Yantai	17	120,000	21,000	-0.83	0.06

5.3.2.2 Testing for Age-Sequential Growth

In Chapter 3 and 4 we analyse the city size distribution in China, furthermore, in last section we analyse the sequential city growth in Chinese cities. In these context, we also inspired by Sanchez-Vidal *et al.* (2014) and seek to test for whether the age of Chinese cities matter in the sequential city growth. In line with Sanchez-Vidal *et al.* (2014), we expect new-born cities to grow faster during the first few years before their growth rate become stable or even declining in the following years.

To test for the age-dependent sequential city growth, we estimate the following model originates from Sanchez-Vidal *et al.* (2014):

$$g_{it} = \alpha + \sum_{k \geq 1} \beta_k \, d_{k,i,t} + \gamma * citysize_{i,t-1} + \delta_t + \theta_r + \varepsilon_{it} \quad (10)$$

Where the explained variable g_{it} is the growth rate of population for each city i at time t (we use yearly data in our analysis rather than decade's data in Sanchez-Vidal *et al.* (2014)) calculated as $g_{it} = lnp_{it} - lnp_{i,t-1}$ (p is the population). $d_k$is a dummy variable which indicates the age of the cities. Specifically, the sub index k in d_k indicates the number of decades that a city existing in our data sample. Thus, when a new city firstly exist in our data sample (our data period is on decade frequency), d_1 (d_k when $k = 1$) equals to one, and zero otherwise. A new city is defined depending on the record in our data sample, if it shows a positive population in one decade and no record previously we define it is a new city. Accordingly, d_2 (d_k when $k = 2$) equals to one indicates that the city existed for one decades in our data sample, and zero otherwise; d_3 (d_k when $k = 3$) equals to one means that the city existed in our data sample for two decades, and zero otherwise, and so on.

Thus, we create the dummy variable d_k to measure the age of a city, from the new-born city (d_1) to eleven decades old (d_{12}). δ_t captures the time fixed effect, θ_r captures the city fixed effect. The variable $citysize_{i,t-1}$ controls for a one decade lag of city size, but, this might cause endogeneity on some level. However, our results show that the impact of a city's age on its growth is robust regardless of including the lagged city size variable or not. ε_{it} is the error term.

5.4 RESULTS

5.4.1 Skewness of City Growth Rate

Firstly we show that the fast-growers are minorities, which is the intuition or foundation of the sequential growth. We find that the distribution of cities' growth rates is skewed to the right in China which is consistent with Cuberes 2011 using Chinese city level data from 1890 to 1994. This shows that the 'fast growers' are a minority as indicated by sequential city growth theory.

In our sample, 94% of the time periods have the positive coefficients of skewness and Figure 5.4- panel A below also shows the density of coefficient of skewness for these periods in our sample. It is apparent from this graph that most of these coefficients are strictly positive. A positive skewness indicates a distribution with an asymmetric tail extending toward more positive values, i.e. few cities are growing fast. Which can be also proved in Figure 5.4- panel B the frequency for the population growth rate of cities for 2010 (0.7 cut-off sample)- most cities' growth rates are concentrated around 0, but a small number of cities grow significantly faster than the others.

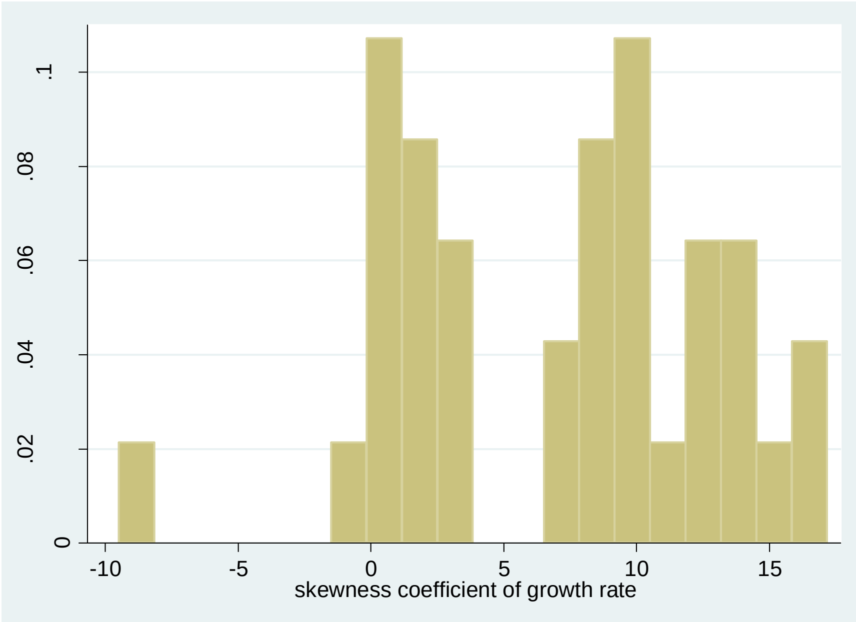


Figure 5.4- panel A: Histogram of density of coefficient of skewness of cities' growth rate from 1890-2010.

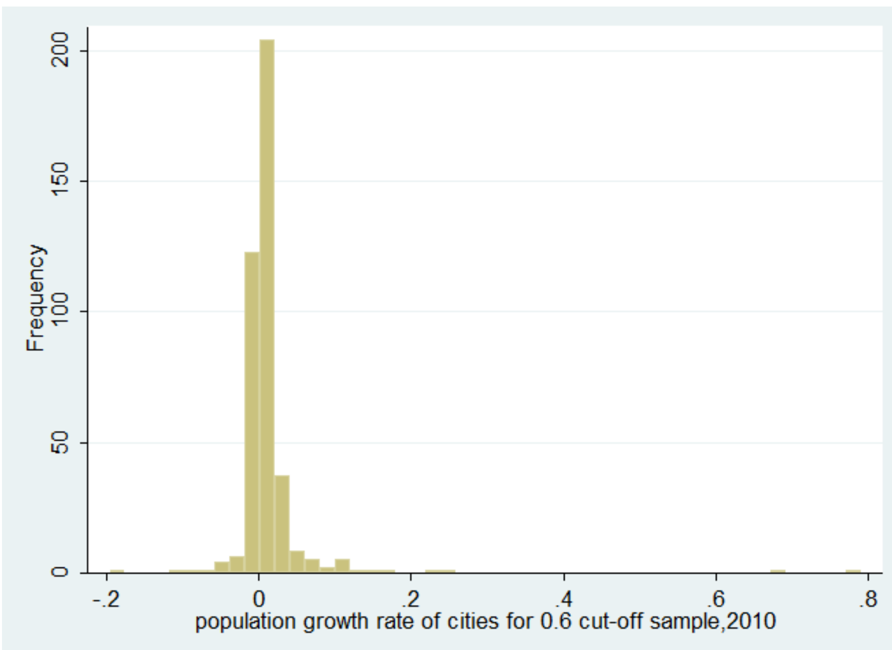


Figure 5.4- panel B: Frequency of population growth rate of cities in 2010 (0.7 cut-off sample)

5.4.2 Average Rank of the Fastest-Growing Cities

In Table 5.3, column [2] shows the number of cities in the 0.7 cut-off sample and column [3] shows the growth rate of increasing number of cities over time. To identify the 'fast-growers' we choose cities whose population growth rate is over ¾ quartile in each year, i.e. the fastest growing 25% of cities ($\boldsymbol{rank_{25}}$). This threshold is reported in column [5] and the number of fast-growers is shown in column [6]. We take the average rank of these fast-growers to get the $\boldsymbol{rank_{25}}$ reported in column [7]. There are year gaps before 1983 due to the data availability, thus the growth rate at that time refers to the decade growth rate.

To be clearer, we show the evolution of $\boldsymbol{rank_{25}}$ over time in Figure 5.5 following Table 5.3. The result is consistent with Cuberes (2011), the average rank of fast-growers is indeed increasing over time from 1890 until the mid-1990s (1994), $\boldsymbol{rank_{25}}$ increases from 4 to 164. This is because the rank of fast-growers' calculated in the $\boldsymbol{rank_{25}}$ is increasing, which might indicate that in the early stage of urbanization large cities are growing the fastest, as time passes the second large cities group (with relatively larger rank) grow the fastest and then the smaller cities become the fastest growers- a sequential city growth. In Cuberes (2011) paper, he uses 54 countries with various time periods panel data to show that within a country city growth follows a sequential growth model. The Chinese data in his paper is from 1890 to 1994. However, we also use the data after 1994 until 2010 and find slightly different results. From mid-1990s to 2001 $\boldsymbol{rank_{25}}$ fluctuates from 151 to 186 and then tends to be stable around 140 after entering the 21st

century, which does not seem consistent with sequential city growth.

More clearly, the trend of $\boldsymbol{rank_{25t}}$ over time is shown in Figure 5.5 below, before 1996 the average rank increases over time especially after 'Economic Reform' (launched in 1979 effect appears around 1983). This may because that the 'Economic Reform' accelerate the process of urbanisation and the more evenly distributed cities (as shown in the above two chaptes), thus promote the relatively small cities to grow faster than before. When entering the 21st century, the average rank of 'fast growers' decreases and is relatively more stable, which implies that the composition of 'fast growers' moves up in the urban size hierarchy and tends to be stable. This trend indicates that basically from the 1890s to the mid-1990s the average rank of the fast-growers is increasing over time, i.e. gradually, the second largest, third largest cities and so on are catching up and growing faster than the initially largest cities in China.

This seems to show empirical evidence for sequential city growth theory- in the early process of urbanization, the largest cities grow fastest. As time passes, population growth in the larger cities declines and the fastest population growth can be found in smaller cities farther down in the urban hierarchy. However, after 2000 the average rank tends to be stable at a rank, one reason would be the number of cities tends to be stable in these years, the other reason might be cities reach a steady state after a sequential growth. In addition, this is plainly in conflict with Gibrat's law in the early years, which claims that city growth rate is independent of its initial size, i.e. larger cities are not growing faster than smaller cities. But after 2000, it is not in conflict with Gibrat's law.

Table 5.3 Time series of average rank of the fastest growing cities (rank25)- 0.7 cut-off sample

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]
Year	no. of cities	growth rate of no. of cities	skewness of city growth rate	growth rate threshold to include into rank25	no. of fast growers	rank 25	Min rank 25	Max rank25	Growth rate of rank25	Urban pop (million)	Urban pop growth rate	Urbanisation rate (%)
1879	31									11.4004		3.11%
1890	8		0.762	41.68%	2	4.000	1	7		6.3366	-44.42%	1.67%
1900	6	-25.00%	1.385	-1.61%	2	5.500	5	6	37.50%	8.1718	28.96%	2.04%
1911	8	33.33%	-0.721	27.19%	2	7.500	7	8	36.36%	9.7625	19.47%	2.28%
1918	10	25.00%	0.91	50.44%	3	4.333	1	10	-42.22%	8.0167	-17.88%	1.74%
1926	12	20.00%	2.642	47.13%	3	8.000	2	12	84.62%	11.6375	45.17%	2.41%
1936	18	50.00%	0.465	85.12%	5	7.000	1	17	-12.50%	23.7842	104.38%	4.68%
1948	28	55.56%	3.064	66.37%	7	15.000	8	23	114.29%	31.373	31.91%	5.73%
1953	44	57.14%	2.381	76.39%	11	23.000	4	35	53.33%	49.7993	58.73%	8.47%
1958	41	-6.82%	2.228	67.02%	11	28.455	10	41	23.72%	65.1313	30.79%	9.96%
1983	93	126.83%	2.236	178.22%	24	50.167	13	90	76.30%	222.74	241.99%	21.62%
1984	87	-6.45%	3.666	6.57%	22	46.045	3	86	-8.22%	240.17	7.83%	23.01%
1985	135	55.17%	8.77	2.49%	34	83.176	15	135	80.64%	250.94	4.48%	23.71%
1986	133	-1.48%	12.289	2.42%	34	74.353	5	132	-10.61%	263.66	5.07%	24.52%
1987	146	9.77%	8.34	2.42%	37	79.676	2	144	7.16%	276.74	4.96%	25.32%
1988	194	32.88%	14.356	2.34%	49	109.184	10	194	37.04%	286.61	3.57%	25.81%
1989	230	18.56%	9.693	2.35%	58	128.897	1	229	18.05%	295.4	3.07%	26.21%
1990	244	6.09%	8.361	2.62%	61	119.246	14	244	-7.49%	301.95	2.22%	26.41%
1991	252	3.28%	-9.488	1.60%	63	138.762	18	252	16.37%	312.03	3.34%	26.94%
1992	260	3.17%	14.935	1.58%	76	147.585	20	260	6.36%	321.75	3.12%	27.46%
1993	289	11.15%	9.409	1.49%	73	138.945	1	289	-5.85%	331.73	3.10%	27.99%
1994	323	11.76%	10.152	1.76%	81	164.568	6	321	18.44%	341.69	3.00%	28.51%
1995	358	10.84%	6.891	1.53%	90	181.022	4	358	10.00%	351.74	2.94%	29.04%
1996	374	4.47%	16.134	1.48%	94	180.564	8	372	-0.25%	373.04	6.06%	30.48%
1999	368	-1.60%	10.261	4.24%	92	158.380	1	366	-12.29%	394.49	5.75%	31.91%
2000	366	-0.54%	6.915	1.94%	92	151.304	3	365	-4.47%	416.08	5.47%	33.35%
2001	376	2.73%	13.277	1.28%	94	186.714	1	374	23.40%	480.64	15.52%	37.66%
2002	360	-4.26%	13.8	1.29%	90	144.633	2	356	-22.54%	502.12	4.47%	39.09%

2003	357	-0.83%	12.354	1.33%	90	145.122	4	352	0.34%	523.76	4.31%	40.53%
2004	352	-1.40%	17.175	1.30%	88	139.966	6	347	-3.55%	542.83	3.64%	41.76%
2005	346	-1.70%	11.778	1.19%	87	137.770	2	346	-1.57%	562.12	3.55%	42.99%
2006	344	-0.58%	0.856	1.49%	86	129.105	1	331	-6.29%	582.88	3.69%	44.34%
2007	339	-1.45%	0.923	1.45%	85	129.482	5	337	0.29%	606.33	4.02%	45.89%
2008	338	-0.29%	9.209	1.22%	85	140.165	3	338	8.25%	624.03	2.92%	46.99%
2009	334	-1.18%	12.289	1.19%	84	136.595	3	333	-2.55%	645.12	3.38%	48.34%
2010	334	0.00%	9.156	1.47%	84	152.506	8	331	11.65%	669.78	3.82%	49.95%
Avg	214.486	14.24%	7.053	19.66%	54.26	99.91	5.66	212.60	15.40%	304.05	17.50%	25.33%

0.7 cut-off refers to the cities in the sample are selected according to the sample selection method in Cuberes (2011). If the city population divided by the mean of the city population in that year is over 0.7, then the city would be selected in the sample. In other words, cities in this sample are over a specific size.
Sources: “China Urban Statistical Yearbook”, *National Bureau of Statistics of China.*
“Population Statistics: historical demography of all countries, their divisions and towns”, *Jan Lahmeyer, http://www.populstat.info/*

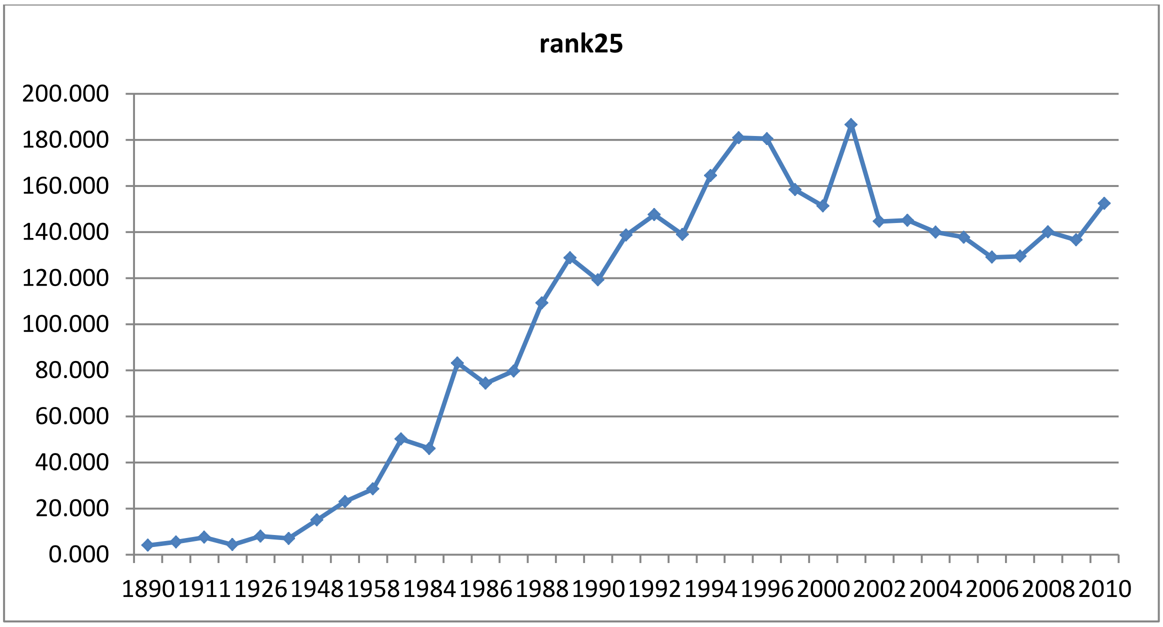


Figure 5.5: Evolution of the average rank of the fastest-growing cities

However, one important concern would be for the reason of the average rank of fast-growers ($\boldsymbol{Rank_{25}}$) increasing over time. We have to be sure that the driving force does not come from the increasing number of cities over time. As showed in Figure 5.6 below, as time passes, not only the average rank of the fast-growers (rank 25) increases, but also the total number of cities in each year in our sample is increasing over time, which can be part of the reason that the average rank of the fast-growers increases.

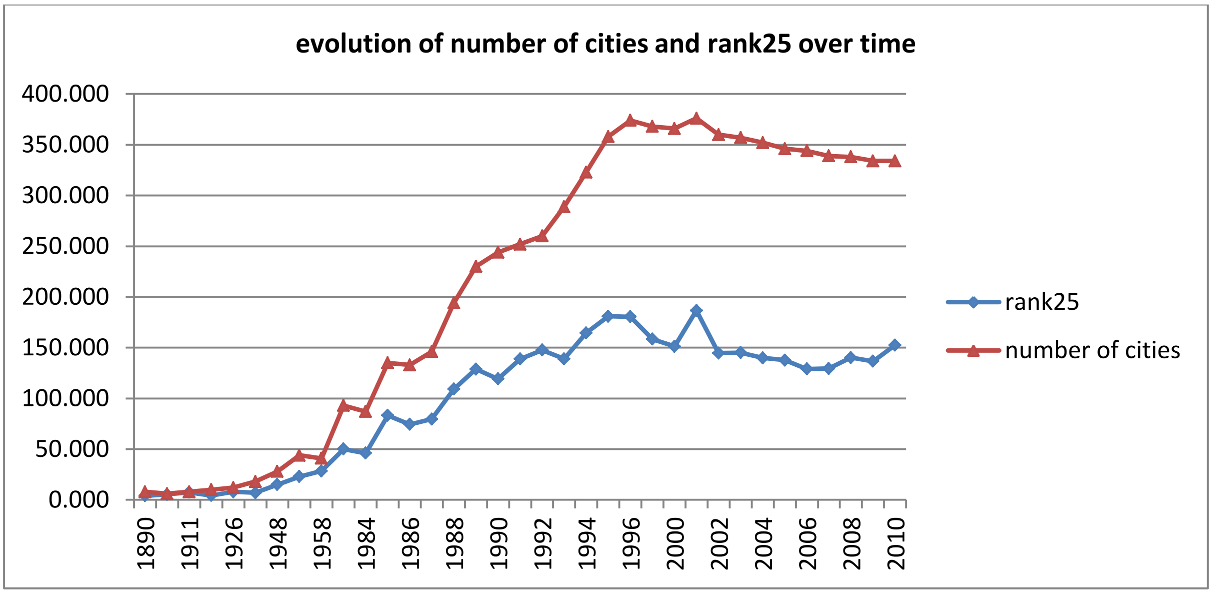


Figure 5.6 Total number of cities in each year

5.4.3 Parametric Analysis for Sequential City Growth

5.4.3.1 $Rank_{25}$ Increases as Time Passes

As Cuberes (2011) suggests, we estimate the equation as we mentioned in methodology section, equation (8), to control the increasing number of cities in each year.

$$\log Rank_{25t} = \beta_0 + \beta_1 t + \beta_2 N_t + \beta_3 N_t^2 + \varepsilon_t \qquad (8)$$

where $Rank_{25t}$ and N_t are the average rank of the top 25% fastest-growing cities and the number of cities in each year, respectively. In addition, we include the square term of the number of cities as a control in order to better capture the relation between the number cities in the sample and the dependent variable. Variable t measures time in years and ε_t is a standard error term. Following Zipf's law literature (Gabaix, 1999) and Cuberes (2011) we use the logarithm of $Rank_{25}$ as the dependent variable to ensure positive predicted values and to minimise the influence of outliers. Results are reported in Table 5.4 below.

Table 5.4: regression of dependent variable on time, the number of cities in each year, and its square term

Dependent var.	Rank25 – $\mathbf{log}\boldsymbol{Rank_{25t}}$		
	(1)	(2)	(3)
Independent var.	lnrank25	lnrank25	lnrank25
Year ($\boldsymbol{t}$)	0.0355***	0.0261***	0.0149***
	(0.00164)	(0.00288)	(0.00289)
Number of cities ($\boldsymbol{N_t}$)		0.00214***	0.0116***
		(0.000575)	(0.00175)
Square number of cities($\boldsymbol{N_t^2}$)			-1.64e-05***
			(2.95e-06)
Constant ($\boldsymbol{\beta_0}$)	-65.82***	-47.89***	-26.48***
	(3.245)	(5.563)	(5.546)
Observations	35	35	35
R-squared	0.934	0.954	0.977

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

We expect that the estimated coefficient for **Year ($\boldsymbol{t}$)** will be positively significant, even after including the **Number of cities ($\boldsymbol{N_t}$)** variable. We find the similar results as shown in Cuberes (2011) for equation (8) that city growth tends to follow a sequential growth model in China. Specifically, in regression (1) of Table 5.4, $\boldsymbol{Rank_{25}}$ is significantly positive correlated with time variable $\boldsymbol{t}$ (at 1% significant level), which implies that the average rank of the fastest-growers (top 25%) in each year tends to increase over time, in other words, every year the faster-growers are including more and more relatively small cities. This is consistent with Cuberes (2009, 2011) who states that early on in the urbanization process the largest cities grow the fastest, as time goes on, the fastest growing cities can be found in smaller cities (rank is relatively larger) and then farther down along the urban hierarchy. This is also consistent with the trend of $\boldsymbol{Rank_{25}}$ over time in Figure 5.5. Next, after taking into account the increasing number of cities over years, the result in regression (2) shows that $\boldsymbol{Rank_{25}}$ still increases as time goes by (there is still a significant positive sign). Only the amount of the time coefficient drops by 20% from 0.0355 to 0.0283. The result shows that after taking into account the increasing number of cities in each year, the average rank of the fastest growing cities is still increasing with time variable, i.e. sequential city growth theory still holds in Chinese cities. The coefficient of the squared term is negative indicating that there might be a concave relationship between $\boldsymbol{Rank_{25}}$ and the number of cities in the sample. It does not notably affect the size of the estimated coefficient of time variable.

To conclude, these results are consistent with the results of Cuberes (2011) studying the combined panel data for a long time period (going as far back as 1790) of 54 countries for cities and 115 countries for metropolitan areas.

5.4.3.2 The Increase in Average Rank is Faster with a Faster Urban Population Growth

Cuberes (2011) also proposed the fact that this sequential city growth process is more pronounced when the urban population grow rapidly. One could test this by identifying whether the growth rate in $\boldsymbol{RANK_{25}}$ is positively correlated with the growth rate in urban population, i.e. by regressing

$$\boldsymbol{g_{RANK_{25t}} = \beta_0 + \beta_1 g_{U_t} + \beta_2 g_{N_t} + \beta_3 Ur_t + \omega_t} \qquad (11)$$

where the $\boldsymbol{g_{RANK_{25t}}}$, $\boldsymbol{g_{U_t}}$, $\boldsymbol{g_{N_t}}$ denotes the growth rate of $RANK_{25t}$, urban population and the number of cities in period t, respectively. $\boldsymbol{Ur_t}$ represents the urbanisation rate. Results show in Table 5.5 below.

Table 5.5: regression of the growth rate of rank25 on the growth rate of urban population and the growth rate of number of cities for each year

Dependent var. growth rate of $RANK_{25}$						
	Zero cut-off			0.7 cut-off		
	(1)	(2)	(3)	(4)	(5)	(6)
Independent var.	grank25	grank25	grank25	grank25	grank25	grank25
Urban pop growth rate	0.442**	0.372***	0.340**	0.377***	0.273*	0.261*
	(0.166)	(0.135)	(0.142)	(0.112)	(0.138)	(0.134)
Number of cities growth rate		0.587***	0.550***		0.324	0.149
		(0.137)	(0.146)		(0.256)	(0.265)
Urbanization rate			-0.357			-0.00667*
			(0.455)			(0.00369)
Constant	0.102	0.0212	0.129	0.0851	0.0588	0.264**
	(0.0801)	(0.0672)	(0.153)	(0.0540)	(0.0574)	(0.126)
Observations	34	34	34	34	34	34
R-squared	0.182	0.486	0.496	0.262	0.298	0.367

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

As seen in Table 5.5, the estimated coefficient on the growth rate of urban population is significantly positive, implying that the $RANK_{25}$ increases faster associated with a more rapid urban population growth, i.e. the middle or smaller cities are more and more becoming the core members in urban population growth, the main force of the growth is farther down the urban hierarchy, when urban growth is more rapid. The result of including the growth of number of the cities shows in regression [2], it weakens the power of growth of urban population on moving down the urban hierarchy (insignificant and lowered value). The growth of number of cities has a significantly positive effect on the $RANK_{25}$ increasing rate, again indicating that the reason of the rank of the fastest-growing cities grows rapidly is not only because of the sequential urban growth, it partially because the cities included in the sample are growing fast.

This result is also consistent with Figure 5.7 below which plots the change of $RANK_{25}$, urban population and the number of cities. In Figure 5.7, when the urban population growth rate increases, i.e. the slope of total urban population increases, the slope of $RANK_{25}$ increases as well, which is consistent with the estimated coefficient of the growth of urban population in equation (11) (significantly and relatively high value of 0.442). Especially after 'China Economic Reform' (launched in 1979 and was effective from 1984), the dramatically faster growth of urbanization correlated with a sharp rise of the average rank of the fast-growing cities, i.e. the rapid growth in China's urban population is associated with a larger slope of $RANK_{25}$. However, only except for the mid 1990s and early 2000s the average rank of fast-growers $\boldsymbol{RANK_{25}}$ decreases a bit while urban population still grows rapidly. This is also in the period of the early stage of the number of cities tends to be stable from mid-1990s to 2010. Thus, the decrease of $\boldsymbol{RANK_{25}}$ might be because of the stable number of cities for each year as time passes, which also indicates the concern we mentioned earlier that the driving force of the rising of $\boldsymbol{RANK_{25}}$ might include the force of the increasing number of cities for each year. To be more precise, we also checked the different time span sample, top 300 cities sample and the fixed 197 cities sample in the next section- robustness check.

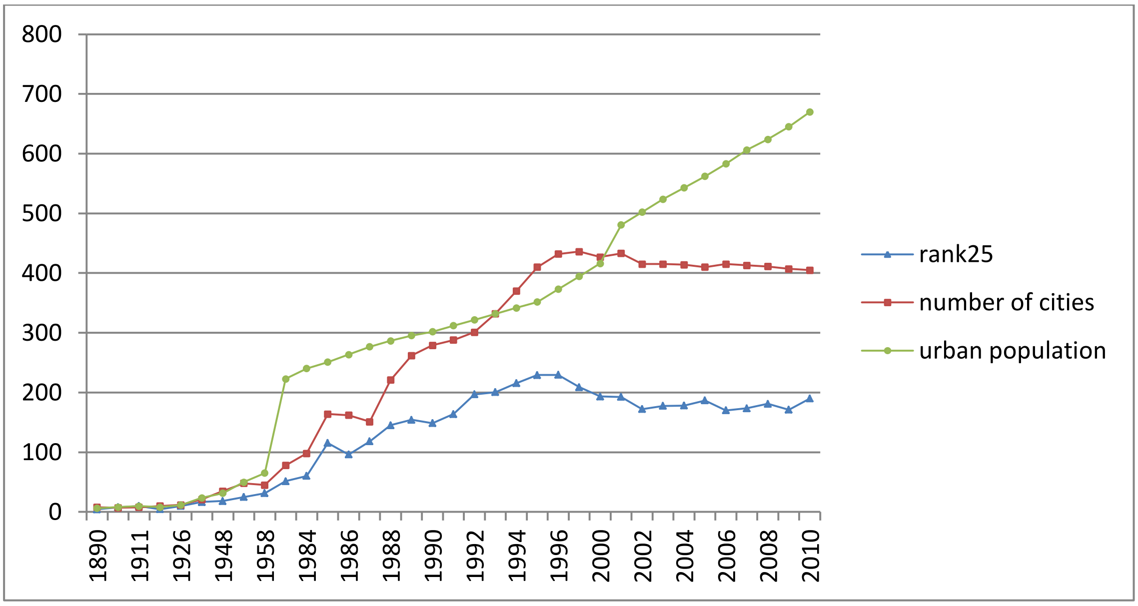


Figure 5.7: Evolution of rank$_{25}$, number of cities and urban population over time (urban population in millions)

5.4.3.3 Additional Subsamples Checks

(1) Different time span samples

Due to the concern of the impact of the increasing number of cities on the result, we also checked different time span samples, a top 300 cities sample (cities involved in the sample may change over year) and a fixed-197 cities sample (cities in the sample are the same over year). Firstly, the different time span samples are basically divided according to the different historical background. We divide the whole sample into four subsamples to attempt to control for the increasing number of cities within the same historical and economic background, as shown in Table 5.6 below: *1)* 1890-2010, from the end of Qing Dynasty to the first decade of the 21st century the number of cities ranges from 8 to 334, the average growth rate of the number of cities during this period is 12.76%; *2)* 1936-2010, from the early second decade of the Republic of China[46] to the first decade of the 21st century, the number of cities ranges from 18 to 334 and the average growth rate is a bit higher at 13.98%; *3)* 1953- 2010, from the first few years of the establishment of the People's Republic of China[47] to the first decade of the 21st century, the number of cities for each year ranges from 44 to 334, the average growth rate is slower at 8.84%; *4)* 1984-2010, from the first 5-year of the Chinese Economic Reform[48] to recent year 2010, the number of cities increases from 135 to 334, the average growth rate of number of cities is much smaller at 5.46%.

[46] Republic of China governed by Nationalist Party was existed from 1927 to 1948 for the mainland China.
[47] People's Republic of China (PRC) governed by Communist Party established in 1949.
[48] Chinese Economic Reform refers to the programme of economic reforms that started in December 1978 by the government. Economic reforms introducing market principles began in 1978 and were carried out in two stages. The first stage, in the late 1970s and early 1980s, involved the decollectivization of agriculture, the opening up of the country to foreign investment, and permission for entrepreneurs to start business. The second stage of reform, in the late 1980s and 1990s, involved the privatization and contracting out of much state-owned industry and the lifting of price controls, protectionist policies, and regulations, although state monopolies in sectors such as banking and petroleum remained.

Table 5.6: Average growth rate of number of cities in each sub-sample and periods

period	Average growth rate of number of cities	Minimum number of cities	Maximum number of cities	Historical background
1890-2010	12.76%	8	336	End of Qing Dynasty -2010
1936-2010	13.98%	18	334	Early second decade of Republic of China
1953-2010	8.84%	44	334	First few year of the establishment of PRC
1984-2010	5.46%	135	334	The first 5-year of Chinese Economic Reform
1995-2010	0.41%	334	358	Migration loosen and Chinese economy changes from 'central planning economy' to 'market economy' until now (*for comparison*)

The parametric results for these subsamples are shown below- we still follow Cuberes (2011) and regress $\boldsymbol{RANK_{25}}$ on time variable:

$$\mathbf{log}\boldsymbol{Rank_{25t} = \beta_0 + \beta_1 t + \beta_2 N_t + \beta_3 N_t^2 + \varepsilon_t}$$

1) The first sample-1890-2010, from the end of Qing Dynasty to the first decade of the 21st century actually is the whole sample that we analysed in the last section in Table 5.3. We find that at an average growth rate of 12.76% for the number of cities increasing in each year, the sequential city growth seems to hold in Chinese cities. *2)* In the 1936-2010 sample, as shown in Table 5.6_A with the growth rate of the number of cities for each year at 13.98%, the estimated coefficient of time variable is significantly positive in regression (1) and still significantly positive in regression (2) after controlling for the number of cities. Results are the same for the adjusted model in regressions (4) and (5). All of these results show the sequential city growth. This also means that the growth of cities which exist from the early stage of the Republic of China until now might follow a sequential city growth model. *3)* For the 1953-2010 sample, during the establishment of the People's Republic of China until now, the average growth rate of the increasing number of cities slows down to 8.84%, and the estimation results are in Table 5.7_B below. Unlike previous estimation results, the estimated coefficient for the time variable is significantly positive, but when we add in the number of cities variable in the regression, the estimated coefficient of time variable become non-significant. This might indicate that the reason for the increasing value of $\boldsymbol{Rank_{25}}$ may be the increasing number of cities in the sample for each year. *4)* Lastly, for the 1984-2010 sample, the average growth rate of the number of cities for each year is quite slow- 5.46%. And the estimation results are a bit contrary to previous ones. For regression (1) in Table 5.7_C, the estimated coefficient for time variable is still significantly positive, but the magnitude has been reduced almost 4 times. And then in regression (2), when we add in the number of cities variable in the regression, the estimated coefficient of time variable become significantly negative, which shows that as time passes the average rank of the fast-growers $\boldsymbol{Rank_{25}}$ is decreasing ($\boldsymbol{Rank_{25}}$ decreasing indicates that fast-growers more and more concentrate to the large cities). This is not consistent with the sequential city growth.

To conclude, the whole sample from 1890 to 2010 (except for the relative $\boldsymbol{Rank_{25}}$ measure) and the 1936-2010 sample show the sequential city growth pattern that large cites tend to grow the fastest first and then the second large cities and so on. However, the sample of 1953-2010 and 1984-2010 seem not support the sequential city growth theory. For the cities experienced the establishment of PRC until now (1953-2010), the increasing $\boldsymbol{Rank_{25}}$ over time maybe because of the increasing number of cities each year. While the cities that experienced the Chinese Economic Reform until now (1984-2010) seem to follow a growth pattern converse to the sequential city growth theory- as time passes the fast-growers will move up to large cities along the urban hierarchy.

Table 5.7_A: 1936-2010 (average number of cities growing rate during this period: 13.98%)

	Rank25 – $\mathbf{log}\boldsymbol{Rank}_{25t}$		
	(1)	(2)	(3)
Explanatory var.	lnrank25	lnrank25	lnrank25
Year ($\boldsymbol{t}$)	0.0427***	0.0275***	0.00660
	(0.00334)	(0.00746)	(0.00408)
Number of cities ($\boldsymbol{N_t}$)		0.00151**	0.0117***
		(0.000673)	(0.00108)
Square number of cities($\boldsymbol{N_t^2}$)			-1.05e-05***
			(1.06e-06)
Constant ($\boldsymbol{\beta_0}$)	-79.59***	-50.13***	-10.58
	(6.649)	(14.55)	(7.889)
Observations	30	30	30
R-squared	0.854	0.877	0.974

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 5.7_B: 1953-2010 (average number of cities growing rate during this period: 8.84%)

	Rank25 – $\mathbf{log}\boldsymbol{Rank}_{25t}$		
	(1)	(2)	(3)
Explanatory var.	lnrank25	lnrank25	lnrank25
Year ($\boldsymbol{t}$)	0.0406***	0.0154	0.00469
	(0.00486)	(0.00950)	(0.00380)
Number of cities ($\boldsymbol{N_t}$)		0.00212***	0.0124***
		(0.000713)	(0.000906)
Square number of cities($\boldsymbol{N_t^2}$)			-1.10e-05***
			(9.27e-07)
Constant ($\boldsymbol{\beta_0}$)	-75.34***	-26.21	-6.933
	(9.691)	(18.61)	(7.417)
Observations	28	28	28
R-squared	0.728	0.799	0.971

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 5.7_C 1984-2010 (average number of cities growing rate during this period: 5.46%)

	Rank25 – $\log Rank_{25t}$		
	(1)	(2)	(3)
Explanatory var.	lnrank25	lnrank25	lnrank25
Year (t)	0.0115***	-0.0166***	-0.0133***
	(0.00384)	(0.00315)	(0.00288)
Number of cities (N_t)		0.00220***	0.00573***
		(0.000213)	(0.00117)
Square number of cities(N_t^2)			-3.61e-06***
			(1.18e-06)
Constant (β_0)	-17.22**	37.60***	30.24***
	(7.670)	(6.192)	(5.766)
Observations	24	24	24
R-squared	0.290	0.883	0.920

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

(2) Top 300 cities sample

To fix the number of cities for each year, we also test the sample with the large top 300 cities each year, the cities included in the sample might change over time. The data descriptive statistics is shown in Table 5.8 below.

Table 5.8 Subsample- Top 300 cities, 1985-2010 (cities might change over time)

Year	no. of cities	no. of matching cities	Threshold of 'fast-growers'	No. of cities classified as 'fast-growers'	skewness of growth rate	rank 25	Min rank of 'fast-growers'	Max rank of 'fast-growers'	Total pop. of top 300 cities	Proportion in total urban pop.	Proportion in national total pop.	urban population (million)	Urbanisation rate (%)
1985	300	274	2.98%	69	9.69	196.71	15	296	210.2949	83.80%	19.87%	250.94	23.71%
1986	300	283	2.77%	71	9.92	197.61	5	298	209.981	79.64%	19.53%	263.66	24.52%
1987	300	245	2.62%	62	9.52	179.31	2	294	248.4258	89.77%	22.73%	276.74	25.32%
1988	300	255	2.46%	64	13.51	179.44	10	299	268.7798	93.78%	24.21%	286.61	25.81%
1989	300	288	2.36%	72	10.17	162.82	1	300	279.7839	94.71%	24.82%	295.4	26.21%
1990	300	289	2.76%	73	8.51	154.74	14	299	291.353	96.49%	25.48%	301.95	26.41%
1991	300	289	1.61%	73	-9.47	165.79	18	300	296.676	95.08%	25.61%	312.03	26.94%
1992	300	270	1.61%	68	14.59	171.03	20	295	312.8071	97.22%	26.70%	321.75	27.46%
1993	300	269	1.49%	68	8.45	136.79	1	299	335.0866	101.01%	28.27%	331.73	27.99%
1994	300	271	1.69%	68	9.55	140.87	6	290	351.8239	102.97%	29.36%	341.69	28.51%
1995	300	289	1.43%	73	9.09	129.01	4	293	361.1268	102.67%	29.82%	351.74	29.04%
1996	300	294	1.44%	74	16.67	131.47	8	299	365.9651	98.10%	29.90%	373.04	30.48%
1999	300	298	4.14%	75	10.46	123.12	1	300	380.2926	96.40%	30.76%	394.49	31.91%
2000	300	297	2.13%	75	5.96	122.51	3	294	396.7068	95.34%	31.80%	416.08	33.35%
2001	300	299	1.30%	75	11.06	127.65	1	299	406.8558	84.65%	31.88%	480.64	37.66%
2002	300	299	1.32%	75	11.74	112.48	2	294	421.0083	83.85%	32.78%	502.12	39.09%
2003	300	300	1.36%	75	10.55	117.31	4	286	428.0715	81.73%	33.13%	523.76	40.53%
2004	300	300	1.41%	75	14.73	119.23	6	300	434.7551	80.09%	33.45%	542.83	41.76%
2005	300	298	1.18%	75	11.49	106.45	2	298	445.4239	79.24%	34.07%	562.12	42.99%
2006	300	298	1.61%	75	-3.60	116.05	1	299	450.092	77.22%	34.24%	582.88	44.34%
2007	300	300	1.53%	75	0.03	119.04	5	290	453.401	74.78%	34.32%	606.33	45.89%
2008	300	300	1.23%	75	8.78	119.28	3	295	457.8795	73.37%	34.48%	624.03	46.99%
2009	300	300	1.21%	75	15.16	121.35	3	299	462.7418	71.73%	34.68%	645.12	48.34%
2010	300	299	1.48%	75	8.90	136.59	8	300	470.7499	70.28%	35.11%	669.78	49.95%
Average	300	287.67	1.88%	72.29	8.98	141.11	5.96	296.5	364.170	87.66%	29.46%	427.394	34.38%

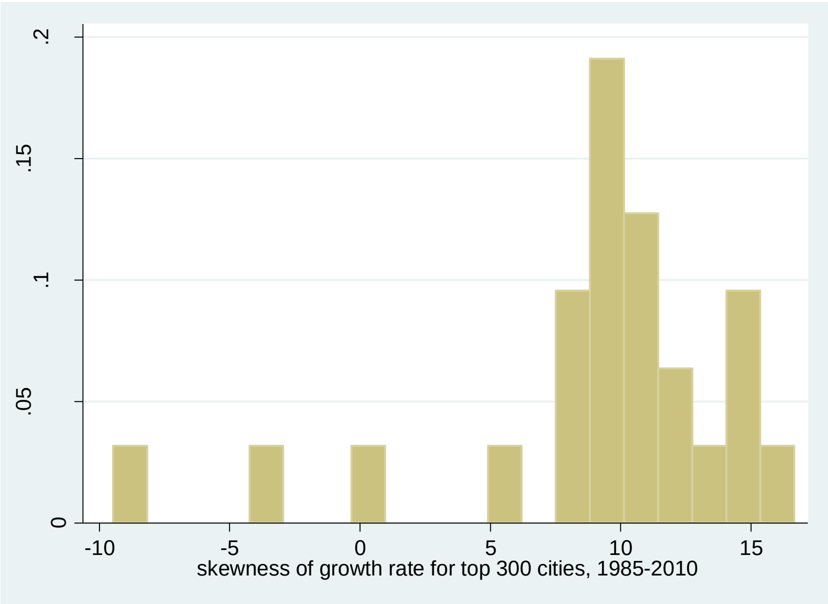


Figure 5.8_A: skewness of growth rate of population for the large top 300 cities

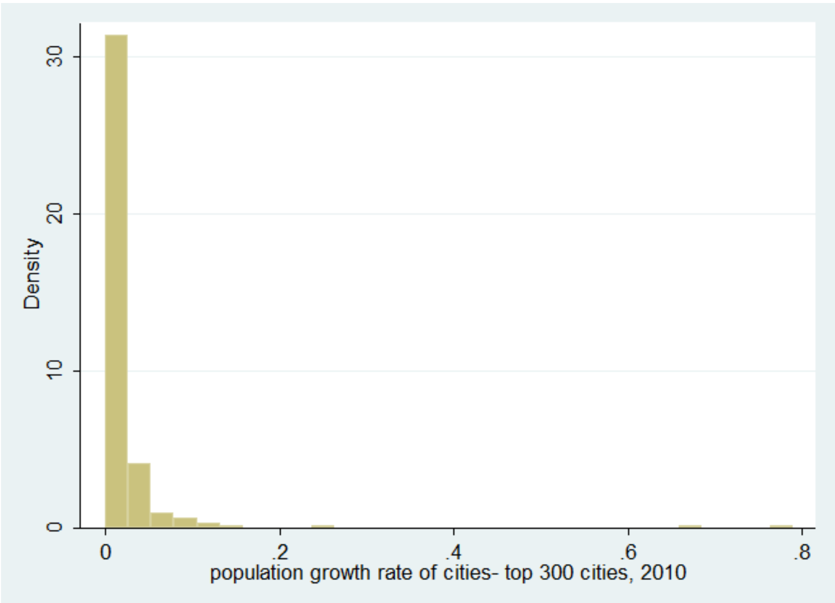


Figure 5.8_B: the density of population growth rate of cities for the large top 300 cities

Firstly, we show the skewness of growth rate of population in Figure 5.8 above, similar to Cuberes (2011) we confirm his result that city growth rate has right skewed distribution to these top 300 cities from 1985 to 2010. In Figure 5.8, coefficients of skewness of city growth rate are mostly concentrated in positive values, i.e. the growth rate of cities is distributed skewed to the right like in Figure 5.8_B. This indicates that during these periods there are only a few cities having the fast growth rate, which is fundamental for seqential city growth.

Secondly, we can observe the trend of the average rank of the fast-growers ($\boldsymbol{Rank_{25}}$) in Figure 5.9 which illustrates Table 5.8. From the graph, it seem that the result is contrary to the sequential theory for the top large 300 cities. The value of $\boldsymbol{Rank_{25}}$ tends to decline from near rank 200 to near rank 100 over time from 1985 to 2005, in the last 5 years to 2010 it increases slightly to rank 150. This indicates that after the Chinese Economic Reform, in the top large 300 cities sample, the relatively small cities (around rank 200) seem to grow the fastest in the 1980s. Then in the 1990s and the first half of 2000s the fast-growers are found in larger cities (city rank increase to around 100), in the last 5 years more smaller cities are categorised into the fast-growers group.

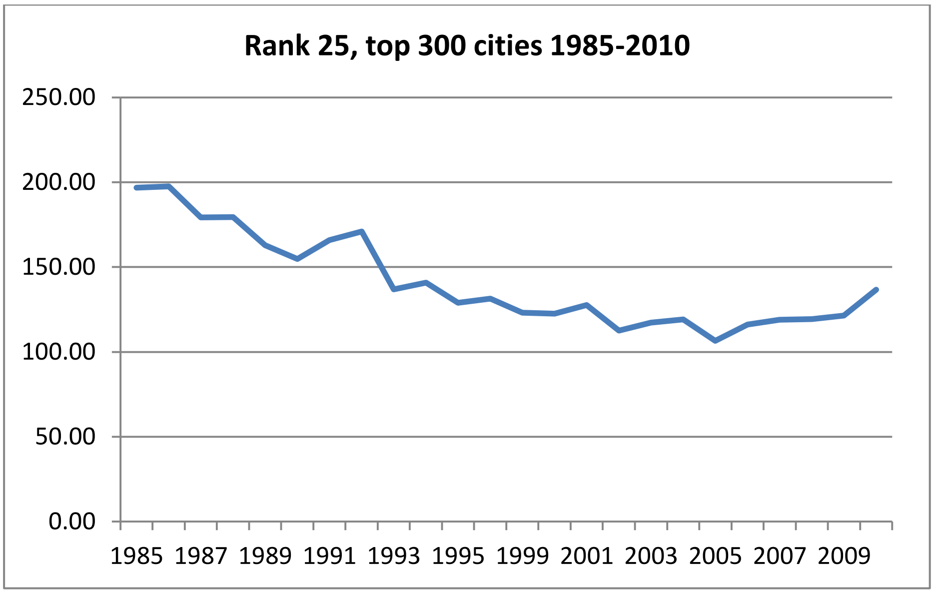


Figure 5.9: average rank of fast-growers over time (made from Table 5.8)

Finally, we show the parametric result using the same regression model as before; Table 5.8 below shows the result. In this sample, the number of cities for each year is fixed, so the estimated coefficient of $\boldsymbol{N_t}$ has been omitted. Consistent with Figure 5.9 above, the estimated coefficient for time variable is significantly negative, which indicates a ‘converse’ sequential city growth pattern. The average rank of 'fast-growers' is decreasing over time, this also shows that the 'fast-growers' are moving up (not down as Cuberes indicates) along the urban hierarchy, i.e. the 'fast-growers' are found more and more in large cities when time passes. The results for $\boldsymbol{Rank_{25}}$and $\boldsymbol{Rank_{25}}/N$ are basically the same (except for the coefficient for constant) because the number of city is fixed. This is consistent with the $\boldsymbol{Rank_{25}}$ trend graph above.

Table 5.9 Top 300 cities, 1985-2010

	Rank25 – top 300 cities sample		
	(1)	(2)	(3)
Explanatory var.	lnrank25	lnrank25	lnrank25
Year ($\boldsymbol{t}$)	-0.0203***	-0.0203***	-0.0203***
	(0.00247)	(0.00247)	(0.00247)
Number of cities ($\boldsymbol{N_t}$)	(omitted)	(omitted)	(omitted)
Square number of cities($\boldsymbol{N_t^2}$)	(omitted)	(omitted)	(omitted)
Constant ($\boldsymbol{\beta_0}$)	45.51***	45.51***	45.51***
	(4.926)	(4.926)	(4.926)
Observations	24	24	24
R-squared	0.755	0.755	0.755

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

(3) Fixed 197 cities sample

As the cities included in the top 300 sample may change over time, we also check the fixed 197 cities sample. From 1985 to 2010 these 197 cities exist all the time and the descriptive statistics is shown in Table 5.9. The method of analysis is the same as for the last sample, firstly we investigate whether there is only a few cities that grow the fastest in each year by checking the skewness of the city growth rate. From Figure 5.10_A and B we confirm that the growth rate for this sample is right skewed, i.e. only a few cities grow the fastest in each year. Secondly, we check the time trend for $\boldsymbol{Rank_{25}}$ in Figure 5.11_A and B which formed from

Table 5.11. Similar to the top 300 cities sample, the average rank of 'fast-growers' firstly declines from 1985 to 2005 and then increases slightly. Lastly, the parametric result shows in Table 5.10, the number of city is fixed, so the estimated coefficient of $\boldsymbol{N_t}$ has been omitted. However, the result is contrary to sequential city size growth pattern as Cuberes (2011) finds. The estimated coefficient for time variable is significantly negative which indicates that the average rank of 'faster-growers' is decreasing over time, similar to top 300 cities sample.

Table 5.10 Subsample- Fixed 197 cities over 1984-2010

Year	no.of cities	threshold-third quartile of growth rate	no.of fast growers	skewness of growth rate	rank25	Min rank of fast growers	Max rank of fast growers	197 cities' total pop (million)	proportion in total urban pop	proportion in national total pop	urban pop (million)	National Total pop (million)
1984	197							160.2	66.70%	15.35%	240.17	1043.57
1985	197	3.01%	50	13.21	134.2	15	197	164.46	65.54%	15.54%	250.94	1058.51
1986	197	2.77%	50	-7.55	127.68	5	196	166.67	63.21%	15.50%	263.66	1075.07
1987	197	2.92%	50	-4.88	122.98	2	196	171.32	61.91%	15.67%	276.74	1093
1988	197	3.11%	50	11.54	129.64	16	193	176.23	61.49%	15.87%	286.61	1110.26
1989	197	2.90%	50	-7.87	130.68	1	196	179.8	60.87%	15.95%	295.4	1127.04
1990	197	2.80%	50	9.34	116.94	14	196	184.75	61.19%	16.16%	301.95	1143.33
1991	197	2.08%	50	-9.48	131.7	18	197	186.48	59.76%	16.10%	312.03	1158.23
1992	197	2.55%	50	3.44	140.32	38	197	189.2	58.80%	16.15%	321.75	1171.71
1993	197	2.57%	50	9.61	133.88	1	197	196.6	59.26%	16.59%	331.73	1185.17
1994	197	3.00%	50	7.27	125.52	20	197	204.17	59.75%	17.04%	341.69	1198.5
1995	197	2.80%	50	5.46	123.96	4	197	211.87	60.24%	17.49%	351.74	1211.21
1996	197	2.68%	50	9.86	122.52	8	197	215.06	57.65%	17.57%	373.04	1223.89
1997	197	2.37%	50	10.84	114.86	1	196	221.96	Fill	in	later	
1998	197	1.97%	50	12.31	113.22	1	197	232.04	Fill	in	later	
1999	197	1.90%	50	-6.39	102.24	1	193	231.69	58.73%	18.74%	394.49	1236.26
2000	197	2.48%	50	12.86	98.98	3	197	239.71	57.61%	19.21%	416.08	1247.61
2001	197	2.28%	50	4.2	87.38	1	196	250.67	52.15%	19.64%	480.64	1276.27
2002	197	2.29%	50	7.97	89.02	2	195	269.86	53.74%	21.01%	502.12	1284.53
2003	197	2.10%	50	8.45	99.12	7	197	279.97	53.45%	21.66%	523.76	1292.27
2004	197	1.93%	50	8.89	86.64	6	197	285.79	52.65%	21.99%	542.83	1299.88
2005	197	1.97%	50	12.35	88.4	4	197	292.5	52.04%	22.37%	562.12	1307.56
2006	197	2.08%	50	4.99	78.72	1	197	300.3	51.52%	22.85%	582.88	1314.48
2007	197	1.94%	50	-0.26	102.78	15	197	304.46	50.21%	23.04%	606.33	1321.29

2008	197	1.48%	50	4.53	92.22	9	197	308.08	49.37%	23.20%	624.03	1328.02
2009	197	1.54%	50	12.13	103.68	8	196	313.63	48.62%	23.50%	645.12	1334.5
2010	197	14.70%	50	6.14	104.18	5	195	320.01	47.78%	23.87%	669.78	1340.91
Average	197	2.85%	50	5.345	111.595	7.923	196.269	231.759	56.97%	18.88%	419.905	1215.323

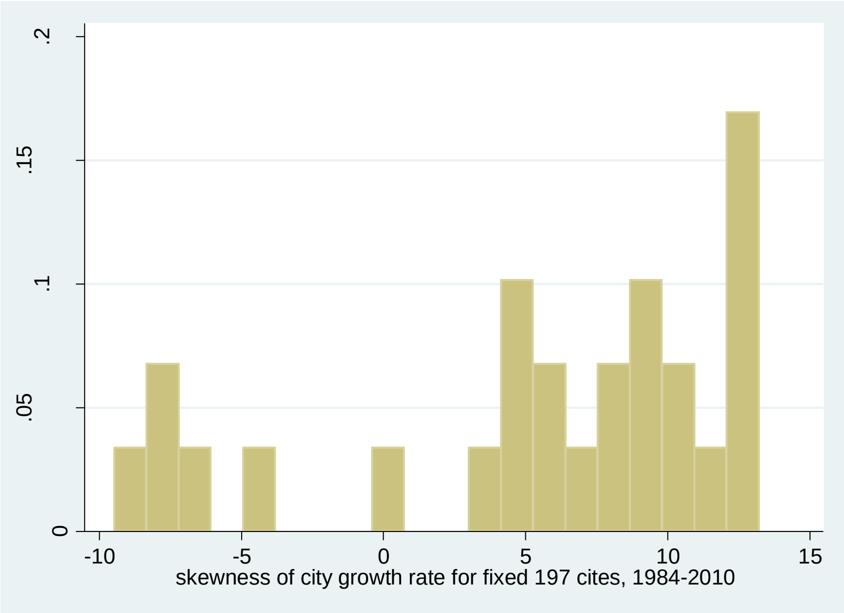


Figure 5.10 average rank of fast-growers over time (made from Table 5.10)

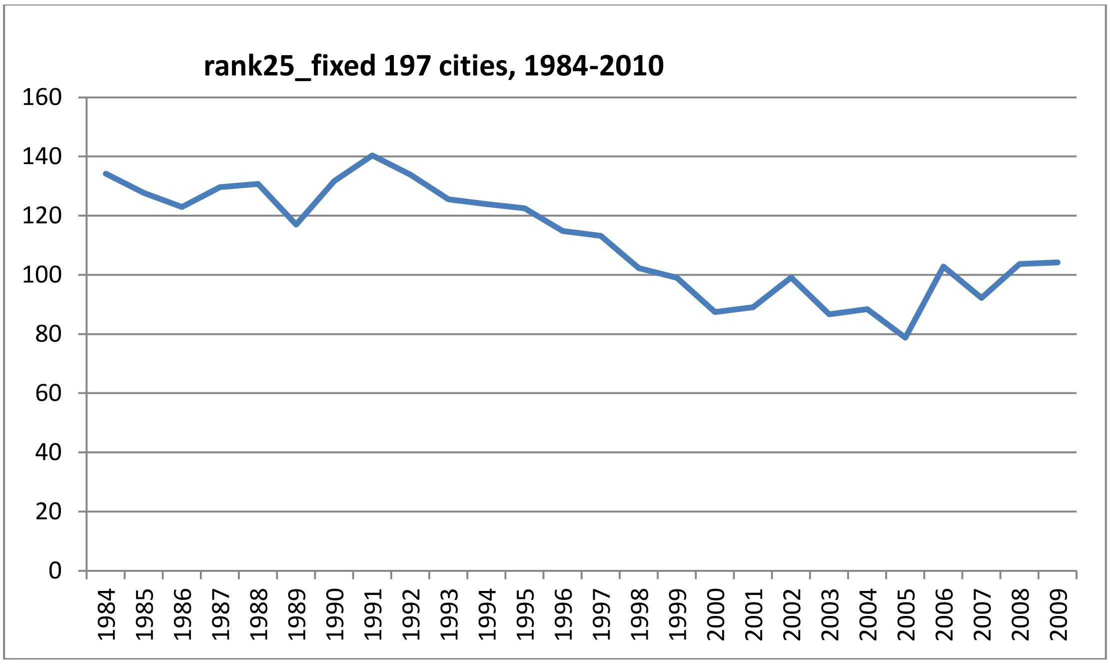


Figure 5.11 average rank of fast-growers over time (made from Table 5.10)

Table 5.11 Subsample- Fixed 197 cities sample, 1984-2010

	Rank25 – fixed 197 cities sample		
	(1)	(2)	(3)
Explanatory var.	lnrank25	lnrank25	lnrank25
Year (t)	-0.0176***	-0.0176***	-0.0176***
	(0.00262)	(0.00262)	(0.00262)
Number of cities (N_t)	(omitted)	(omitted)	(omitted)
Square number of cities(N_t^2)	(omitted)	(omitted)	(omitted)
Constant (β_0)	39.90***	39.90***	39.90***
	(5.230)	(5.230)	(5.230)
Observations	26	26	26
R-squared	0.654	0.654	0.654

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

To conclude, according to Cuberes (2011), firstly we confirm that the growth rate of cities is right skewed in China over the last century (1890-2010) as many other countries tested in Cuberes (2011), which indicates that only a few cities grow the fastest in each year. Secondly, sequential city growth can be found in Chinese cities over 130 years (1879 to 2009) and accordingly this paper documents two novel empirical facts about Chinese city growth. The first is that the average rank of the fast-growing cities increases over time, which indicates the sequential urban growth pattern, i.e. the largest cities are the first to grow fast, as time goes by, the fastest grower can be found in middle and then smaller cities, farther down along the urban hierarchy. Secondly, empirical results show that this sequential city growth pattern in China is even more noticeable when the urban population is growing rapidly, like since the 'Economic Reform' (1984) the process of urbanization dramatically increases and the average rank of fast-growers increases sharply.

Next, as the concern of the driving forces of the increasing average rank of the 'fast-growers' might include the increasing number of cities in each year, although we controlled for this in the parametric estimation by adding the number of cities variable and its square term, we also test for some subsamples trying to control for this factor.

Firstly we test different time spans sample according to the historical background, because the growth rate of number of cities in these different time spans are slowing down from 12.76% to 5.46%. For the first two samples *1)* 1890-2010 (actually, it is the same sample as above, including here for historical background completeness) and *2)* 1936-2010 sample shows the sequential city growth pattern; however, the last two samples *3)* 1953-2010 and *4)* 1985-2010 do not show sequential city growth, the former sample had an insignificant coefficient for the time variable and the latter sample had a negative coefficient for the time variable, which is contrary to the sequential city growth. This might be because the sequential city growth will show in the long run, for the last two samples there are only 5 decades and 2 decades time for cities to grow.

Additionally, in order to control the increasing number of cities for each year as time passes, we test for the top large 300 cities sample and fixed 197 cities from 1985 to 2005 and find that the estimated coefficient for time variable are all significantly negative which means that the average rank of the 'fast-growers' decrease as time passes, i.e. relatively small cities tend to grow the fastest first and then as time passes large cities tend to grow the fastest in these samples. It seems like sort of converse sequential growth. However, we have to notice that the results are from the sample within relatively large cities (the fixed 197 cities from 1985 to 2010 also are relatively large cities as they exist during the whole period), we might conclude that within large cities, in the last almost three decades, there may show a converse sequential city growth.

5.4.4 Age Sequential

As mentioned in last size sequential city growth section, the number of cities each year is growing rapidly from 18 (in 1890) to 654 (in 2010) over the last two centuries. The including of new cities in the size sequential model may bias the method used to verify the theory, as the increasing of average rank of 'fast-growers' over time may because of the increasing number of cities each year. Thus, we can also test whether cities grow relevant to their age, as recently Sánchez-Vidal *et al.*(2014) did with U.S. city level data from 1900 to 2000.

Figure 5.12 below depicts the evolution of the total number of Chinese cities over time, throughout the twentieth century. From this graph we can observe that the number of cities increases over time. Specifically, the growth of number of cities is stable during 1890s and 1920s as agriculture was dominating the economic growth; from mid of 1920s to mid of 1950s, the number of cities grew gradually as the establishment of the Republic of China [49]and People's Republic of China[50]; there is a bit decrease from mid 1950s to mid-1960s due to the institutional reorganise as a new country, and quite stable from mid-1960s to mid-1970s because of the 'Culture Revolution' [51]; a sharp increase was occurred after the 'Economic Reform' [52]from late 1970s until early 21st century, the number of cities grew from around 100 to over 600, and then stay stable until now.

Table 5.12 shows the descriptive statistics for the number of cities and population size in each decade from 1890 to 2010. One can observe that the number of cities increases over, except for the 1950s and 1960s as these two decades experienced the establishment of a new

[49] Republic of China is from 1912 to 1949.
[50] People's Republic of China is from 1949 until now.
[51] 'Culture Revolution' is from 1966 to 1976.
[52] 'Economic Reform' was launched in 1979.

government (1959) thus the number of cities may have fluctuated a lot. The number of cities in 2010 is almost 36 times that of 1890, which indicates the importance of taking the new cities into account when we study China's urban growth process. Table 5.12 illustrates the urbanisation process of China over the last century.

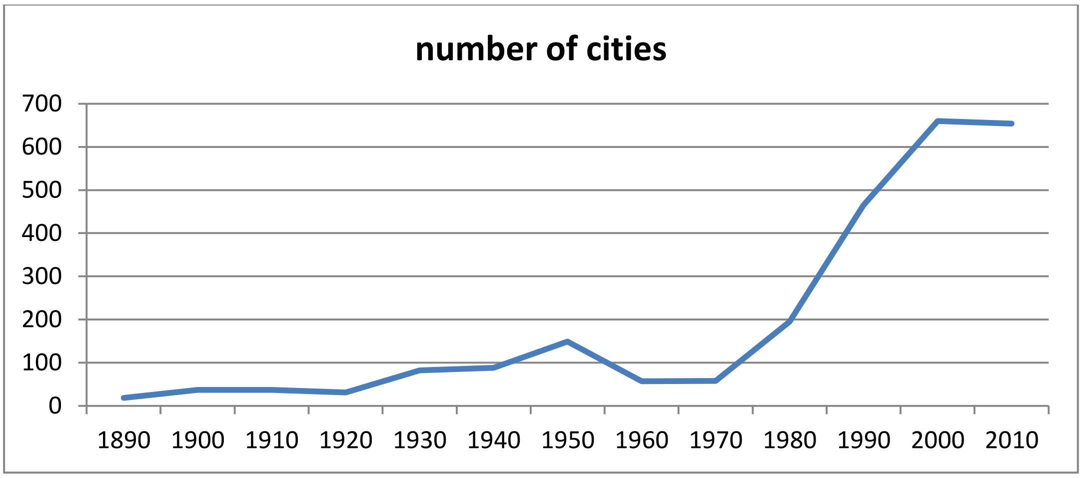


Figure 5.12: Number of cities growing over time

Table 5.12 descriptive statistics for cities

year	number of cities	mean size	standard deviation	minimum	maximum
1890	18	352,033.33	447258.9	3700	1600000
1900	37	261,378.38	266717.8	20000	900000
1910	37	282,300.00	297758.1	20000	1135000
1920	31	343,432.26	373171.6	18500	1498633
1930	82	287,033.33	442389.7	32000	3403000
1940	88	356,511.36	566471.9	7000	4423000
1950	149	380,446.98	726910.3	30000	6590700
1960	57	1,089,666.67	1207051	77000	6431000
1970	58	1,121,250.00	1103371	85000	5686000
1980	195	669,107.09	827651.8	73700	6628367
1990	465	720,313.12	732412.1	9600	7834800
2000	660	544,255.22	952597.6	15000	1.14E+07
2010	654	970,004.88	1198628	48000	1.54E+07

Table 5.13 below shows the value of the age dummies over the last century. As mentioned in

the methodology section, d_1 (d_k when $k = 1$) represents the number of new cities emerged in that decade, i.e. in 1900, a total number of 17 cities emerged; in 1910, the number was 20, and so on. Similarly, d_2 (d_k when $k = 2$) is the number of cities existed for one decades in our data, i.e. in 1910 there were 17 cities existing for one decades; in 1920 there were 10 cities existing for one decades. Accordingly, column d_3 indicates the number of cities existing for two decades (20 years old in our data sample), and similarly until d_{12}(cities existing for 110 years throughout our data sample). The total number of cities for each decade shows in the last column of Table 5.13.

Furthermore, with this table (Table 5.13) one can trace the evolution of the cities from the first decade they emerge until the end by observing the diagonals. Specifically, as d_1 indicates the number of new cities per decade; and d_2 shows the number of cities existing for two decades; d_3 are those with three decades existence each decade and so on, this enables us to trace the 17 new-born cities in 1900 by observing the number of cities in d_2 in 1910, d_3 in 1920, etc.. According to this construction, the numbers of cities in diagonal in Table 5.13 will not increase over time. However, we can observe that some cities may disappear over time, as the number in the diagonals is decreasing gradually. This might because that during this long period, some cities might expand or absorb others. Furthermore, this decreasing trend was obvious between 1950 and 1960, because the People's Republic of China was established during this period and the city's definition and system may have been changed significantly.

Table 5.13: evolution of new cities 1900-2010

year	d1	d2	d3	d4	d5	d6	d7	d8	d9	d10	d11	d12	total
1900	17	0	0	0	0	0	0	0	0	0	0	0	17
1910	20	17	0	0	0	0	0	0	0	0	0	0	37
1920	4	10	17	0	0	0	0	0	0	0	0	0	31
1930	58	1	8	15	0	0	0	0	0	0	0	0	82
1940	6	58	1	8	15	0	0	0	0	0	0	0	88
1950	61	6	58	1	8	15	0	0	0	0	0	0	149
1960	4	3	1	26	1	7	15	0	0	0	0	0	57
1970	0	4	3	1	26	1	8	15	0	0	0	0	58
1980	139	0	4	3	1	26	1	6	15	0	0	0	195
1990	277	134	0	4	2	1	26	1	6	14	0	0	465
2000	210	265	132	0	4	2	1	26	1	6	14	0	661
2010	24	190	258	131	0	4	1	1	26	1	6	14	656
total	820	688	482	189	57	56	52	49	48	21	20	14	1545

Created by our dataset.

As shown in the methodology section, we analyse the age-sequential growth for cities by estimating the following model mentioned in the methodology section as equation (10) according to Sánchez-Vidal *et al.*(2014):

$$g_{it} = \alpha + \sum_{k\geq 1} \beta_k \, d_{k,i,t} + \gamma * citysize_{i,t-1} + \delta_t + \theta_r + \varepsilon_{it} \quad (10)$$

where g_{it} is the growth rate for each city i at time t, in our case the time period is at decade frequency. g_{it} is calculated as $g_{it} = lnp_{it} - lnp_{i,t-1}$, where p being the population. As mentioned before, d_k is a dummy variable capturing the age of the cities. δ_t represents the time fixed effect, θ_r represents the city fixed effect. The variable $citysize_{i,t-1}$ controls for a one decade lag of city size. ε_{it} is the error term.

According to the age sequential growth hypothesis (Sánchez-Vidal *et al.*, 2014), that young cities or new-born cities grew the fastest initially, β_k is expect to be positive and significant during the first decades following the birth of a city, and then as time passes, we would expect this estimated coefficient to decrease or even showing a negative value. In addition, we use fixed effect model to control for the time-invariant unobservable influences within groups.

Table 5.14 shows results of estimating for equation (10). Implicitly, we omit d_0 in all of the estimations, as d_1 represents cities existing for one decade which is calculated from d_0. The estimated coefficients interprets the average impact of city age on the growth rate of a specific city i, with respect to the average growth rate of the whole sample. As explained above, d_1 indicates cities when it was newly born (have positive record in the current decade, but no record in previous period), thus, d_2 indicates cities existing for one decade; d_3 two decades and so on, so that d_{12} represents cities with 110 years old.

Therefore, the estimated coefficient of d_1 can be interpreted as the additional average impact on city growth of being a new-born city, with respect to the average growth of the whole sample from 1890 to 2010; estimated coefficient associated with d_2 represents the additional average impact of one-decade years old cities on the city growth rate, and so on. Thus, we tend to be more interested in the trend of the estimated coefficients from d_1 to d_{12}, which indicates the dynamic impacts of a city's age on their growth rate with respect to the average growth rate for the whole sample. Due to the long period we studied, the average growth rate interprets the long run growth rate, and the estimated coefficients of our dummies can show the dynamics of city growth.

Table 5.14 below shows the estimated regression results for the impacts of city's age on growth. We expect that the estimated age coefficients are significant and the magnitude of the coefficients show a decrease trend from new-born to mature cities if city growth is consistent with the studies by Sánchez-Vidal *et al.*(2014). Regression (1) shows the results of estimating equation (10) by OLS without any control variables. Regression (2) shows the same estimation but including the one period-lag of city size as a control variable. The estimated coefficients in regression (1) are not significant, but when we control for city size in Regression (2) the estimated coefficients for 11 age dummies become significant but increasing, which indicates that older cities were growing faster, as time passes, younger cities grow more slowly in China. This is contrary to the results of Sanchez-Vidal *et al.* (2014) showing that new-born cities grow the fastest in U.S., and this may be explained that firstly, in China, historical cities have long been the attracted migration area as they may have better amenities and better job opportunities and mature urban system and infrastructure than new cities; secondly, urban policy tends to favor historical and large cities in China during last century.

In addition, as we study the age impacts on city growth, there might a considerable amount of uncontrolled information missing in Regression (1) and (2) using OLS estimation. In order to control those possible biases, we estimate regression (3) using the fixed effect model to control the time-invariant unobservable factors that might simultaneously affect the LHS and RHS of the regression. Results are similar to regression (2), only the magnitudes are greater. The estimated parameters show how new-born city i grows in decade $t > 1$ in comparison with how new-born city i grew in decade$t$. The trend of the estimated coefficients from d_1 to d_{11} is decreasing, which indicates the inverse-age sequential city growth in contrast to U.S. cities (Sanchez-Vidal *et al.*, 2014).

The results might indicate that in China, the growth rate of a new-born city is smaller than the growth rate when it becomes mature. Specifically, in China, the average age impact on city growth in the first decade since it has been created (10 years old city) is about 3.874 points lower than the average growth rate of the whole period (which corresponds to the estimated value of the constant in the regression, 18.12). One decade later, the age impact becomes a bit stronger, as the estimated coefficient falls to 3.182 lower than the average growth rate for the whole period. As the age of the city increases, the age impact on the growth is stronger. Thus, the higher growth occurs during the later decade of a city's existence in China.

In addition, we also estimate regression (4) to regression (8) which reports the result for the prefecture-level cities, East region cities, Middle region cities, West region cities and Northeast region cities respectively. For regression (4), prefecture level city ranks below a province and above a county level city in China's administrative structure. In principle it represents all the main cities in China. For this group, the trend of the estimated coefficients is similar to previous estimations; new cities grow slower while when they become mature they tend to grow faster than before. Nevertheless, the overall size of the coefficients is smaller than before, which might indicate that the impact of the variation of age on the variation of the growth rate in prefecture level cities is small, i.e. they tend to grow in a more stable manner among different aged cities compared to other county level cities. Regression (5) shows the result for the East region cities, all the estimated coefficients are almost the same as the fixed effect model- regression (3), either in terms of significance or the sign or the magnitude. This is consistent with the fact that the East region cities form the majority of the fast growing cities. This might be because the economic development always occurs and keeps developed in the East region, from end of Qing dynasty to the modern time. Regression (6) considers for the Middle region cities and the result is still similar to the fixed effect regression (3), Prefecture level cities regression (4) and the East region cities regression (5). The only difference is that the size of the coefficients is a bit smaller than the East group, which reveals that the impact of age on the variation in growth rate for Middle region cities is smaller than East cities. Regression (7) for the west region cities, result is consistent with previous ones, showing the inverse-age sequential growth. Lastly, the result for the Northeast region cities in the last regression (8) is a bit different to all the previous regressions. The estimated coefficients are all positively significant and the magnitudes show an increased trend, which reveals that, for the Northeast region cities, the more mature city the faster its growth. This is also contrary to the age sequential theory, which might be because the Northeast region [53]has long been the major heavy industrial bases of China, thus the mature cities continuously receive favourable development policies to support high economic growth.

To conclude, the city growth pattern with respect to cities' age in China is contrary to what has been found in the U.S. cities by Sanchez-Vidal *et al.* (2014). We find that within a decade, older cities grow faster, while new-born cities grow relatively slower. However, in the following decades the growth rate of young cities accelerate a bit and the new-born cities in each decade show slower growth rate than other cities. Thus, an inverse-age sequential city growth pattern is found in Chinese cities from 1900 to 2010.

[53] The Northeast region consists of the three provinces of Liaoning, Jilin and Heilongjiang, also known as the Three North-eastern Provinces.

Table 5.14 estimation for age sequential

Dependent variable: population growth at the city level								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Decades of existence	OLS	OLS	FE	Prefecture	East	Middle	West	Northeast
d1	-0.168	-2.445***	-3.874***	-1.591***	-3.773***	-2.712***	-4.749***	
	(0.283)	(0.225)	(0.333)	(0.220)	(0.417)	(0.422)	(0.626)	
d2	-0.133	-1.896***	-3.182***	-1.293***	-3.540***	-2.295***	-3.887***	1.774**
	(0.246)	(0.197)	(0.254)	(0.204)	(0.321)	(0.300)	(0.455)	(0.686)
d3	0.322	-1.427***	-2.576***	-1.153***	-2.786***	-1.634***	-3.536***	2.236***
	(0.248)	(0.198)	(0.263)	(0.194)	(0.357)	(0.326)	(0.438)	(0.649)
d4	-0.00387	-1.338***	-2.366***	-1.003***	-2.560***	-1.464***	-3.210***	2.446***
	(0.253)	(0.198)	(0.237)	(0.175)	(0.296)	(0.282)	(0.414)	(0.610)
d5	-0.181	-1.340***	-2.094***	-0.994***	-2.406***	-1.584***	-2.770***	3.066***
	(0.289)	(0.226)	(0.213)	(0.156)	(0.263)	(0.265)	(0.359)	(0.575)
d6	-0.118	-1.179***	-1.842***	-0.846***	-2.042***	-1.353***	-2.510***	3.276***
	(0.282)	(0.220)	(0.210)	(0.155)	(0.279)	(0.232)	(0.376)	(0.569)
d7	-0.0900	-1.016***	-1.617***	-0.725***	-1.840***	-1.192***	-2.207***	3.621***
	(0.275)	(0.211)	(0.204)	(0.146)	(0.265)	(0.194)	(0.335)	(0.594)
d8	-0.182	-0.906***	-1.375***	-0.680***	-1.526***	-0.844***	-2.056***	3.797***
	(0.276)	(0.211)	(0.182)	(0.127)	(0.226)	(0.201)	(0.260)	(0.622)
d9	-0.121	-0.724***	-1.119***	-0.544***	-1.178***	-0.661**	-1.776***	3.930***
	(0.277)	(0.211)	(0.176)	(0.126)	(0.205)	(0.278)	(0.238)	(0.642)
d10	-0.0229	-0.576**	-0.872***	-0.365***	-0.974***	-0.426*	-1.503***	4.462***
	(0.315)	(0.236)	(0.151)	(0.129)	(0.168)	(0.217)	(0.261)	(0.758)
d11	0.0618	-0.271	-0.409***	-0.126	-0.480***	0.196***	-0.828***	4.588***
	(0.318)	(0.238)	(0.128)	(0.124)	(0.130)	(0.0611)	(0.292)	(0.822)
o.d12	-	-	-	-	-	-	-	
lnL_10pop		-0.716***	-1.178***	-0.475***	-1.307***	-1.213***	-1.131***	-1.177***
		(0.0219)	(0.0705)	(0.0618)	(0.0908)	(0.128)	(0.154)	(0.165)
o.pref				-	-	-	-	-
o.d1								-
d12								5.392***
								(0.894)
Constant	0.348	11.12***	18.12***	7.564***	20.13***	17.74***	18.14***	13.15***
	(0.244)	(0.379)	(1.068)	(0.952)	(1.393)	(1.906)	(2.373)	(1.697)
Observations	1,545	1,545	1,545	947	582	394	366	218
R-squared			0.733	0.289	0.844	0.741	0.660	0.704
Number of cities	644	644	644	267	252	166	148	88

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

5.5 CONCLUSION

In this paper, we study the evolution of city growth in China using city level data from 1879 to 2009. We document two novel empirical facts. The first is about sequential city growth, we show that city growth rates are clearly skewed to the right each year in China, indicating that there are a few cities that grow much faster than the rest in each year. Thus, furthermore, we investigate that the rank of these fast growing cities rises as time goes by, implying that early on in the process of urbanization, fast-growers are concentrated in large-sized cities initially and then as time passes, the fast-growers can be found in middle-sized cities and then relatively small cities and so on. In other words, cities grow in sequential order, with the initially largest ones being the first to develop.

In addition, this sequential growth process is more pronounced where urban population grows rapidly. In addition, we test for different subsamples attempting to test the validity of this sequential growth theory with slower growth rate of number of cities increasing over time, or even without the number of cities growing over time.

Results for different subsamples are shown in Table 5.15 below. We can see from Table 5.15 for the first two samples, cities show sequential growth, however, after the establishment of PRC cities do not show sequential growth pattern, or converse sequential (small-sized cities grow faster initially, and then large-sized cities). The fixed number of cities sample in the last two rows represents the large cities (top 300) and the cities exist for almost the last three decades, they tend to follow converse sequential city growth, small cities may be the first to grow and then the fast-growers move to the large cities.

Table 5.15 Sequential city growth results table

Sample year	Sample cities	Whether sequential growth	Average growth rate of number of cities	Historical background
1890-2010	Range from 8 to 334	Yes	12.76%	End of Qing Dynasty -2010
1936-2010	Range from 18 to 334	Yes	13.98%	Early second decade of Republic of China
1953-2010	Range from 44 to 334	Not shown	8.84%	First few year of the establishment of PRC
1984-2010	Range from 135 to 334	converse	5.46%	The first 5-year of Chinese Economic Reform
1985-2010	Top 300 cities	converse	0%	
1985-2010	Fixed 197 cities	converse	0%	

Our second finding relates to the impact of the age of cities on the evolution of city growth. We have performed our analysis focusing on the role played by the new-born cities that have been created during the decades of our analysis. Using parametric analysis we obtain the result that there are differences in city growth rates according to the age of the city. In general, when a city is born, it has the lowest growth rate; however, as decades pass, when it matures its growth rate accelerates. This result is contrary to Sanchez-Vidal *et al.* (2014). The trend of the coefficient is that the values are increasing from young cities to older ones, which means even though the growth rate for all cities are not very fast but older cities still show faster growth than younger ones, thus the inverse-age sequential city growth may hold in China. it may be because that firstly, historical cities have long been the attracted migration area as they may have better amenities and better job opportunities than new cities; secondly, urban policy favor historical and large cities in China. We also test for the prefecture-level cities and four different economic regions. Results are basically the same as the whole sample, except for the north-eastern region. In the north-eastern region, the impact of the age of a city on the growth of a city is greater than other groups. This might be because the north-eastern region has long been the heavy industrial base and the more mature cities receive the more favourable policies and thus grow faster.

APPENDIX

Table A5.1: -Whole sample- zero cut-off

Year	no. of cities	no. of cities growth rate	no. of matching cities	Skewness of growth rate	rank25	rank25/no. of cities	Growth rate rank25	grate of rank25/no	Min rank25	Max rank25	Urban pop	Urban pop growth rate	Urbanisation rate	National total pop
1879	31		16								11.4004		3.11%	366.988
1890	19	-38.71%	16	0.2279	7	0.368			1	13	6.3366	-44.42%	1.67%	380
1900	25	31.58%	17	4.0293	18	0.720	157.14%	95.43%	9	22	8.1718	28.96%	2.04%	400
1911	38	52.00%	23	1.1755	17.17	0.452	-4.61%	-37.24%	5	30	9.7625	19.47%	2.28%	427.662
1918	30	-21.05%	17	0.5463	8.8	0.293	-48.75%	-35.08%	1	21	8.0167	-17.88%	1.74%	461.766
1926	31	3.33%	27	2.7808	19.57	0.631	122.39%	115.21%	13	30	11.6375	45.17%	2.41%	482.128
1936	82	164.52%	24	0.4196	16.83	0.205	-14.00%	-67.49%	1	33	23.7842	104.38%	4.68%	507.864
1948	88	7.32%	82	3.3855	37.05	0.421	120.14%	105.13%	8	77	31.373	31.91%	5.73%	547.804
1953	153	73.86%	88	6.7856	45.24	0.296	22.11%	-29.77%	4	89	49.7993	58.73%	8.47%	587.96
1958	127	-16.99%	92	3.1034	32.83	0.259	-27.43%	-12.57%	2	81	65.1313	30.79%	9.96%	654.159
1983	195	53.54%	116	2.0261	69.41	0.356	111.42%	37.70%	10	181	222.74	241.99%	21.62%	1030.08
1984	207	6.15%	187	3.6896	114.98	0.555	65.65%	56.05%	3	204	240.17	7.83%	23.01%	1043.57
1985	324	56.52%	294	10.0169	213.22	0.658	85.44%	18.48%	15	324	250.94	4.48%	23.71%	1058.51
1986	321	-0.93%	301	10.1224	205.09	0.639	-3.81%	-2.91%	5	321	263.66	5.07%	24.52%	1075.07
1987	382	19.00%	318	6.9744	243.91	0.639	18.93%	-0.06%	2	382	276.74	4.96%	25.32%	1093
1988	434	13.61%	378	12.5915	280.38	0.646	14.95%	1.18%	16	431	286.61	3.57%	25.81%	1110.26
1989	449	3.46%	430	12.2407	281.61	0.627	0.44%	-2.92%	1	449	295.4	3.07%	26.21%	1127.04
1990	467	4.01%	446	9.9770	259.35	0.555	-7.90%	-11.45%	14	466	301.95	2.22%	26.41%	1143.33
1991	478	2.36%	464	-9.8751	306.91	0.642	18.34%	15.61%	18	475	312.03	3.34%	26.94%	1158.23
1992	517	8.16%	475	17.4947	334.81	0.648	9.09%	0.86%	23	516	321.75	3.12%	27.46%	1171.71
1993	570	10.25%	465	11.5709	368.45	0.646	10.05%	-0.18%	1	569	331.73	3.10%	27.99%	1185.17
1994	620	8.77%	563	12.3405	374.04	0.603	1.52%	-6.67%	6	619	341.69	3.00%	28.51%	1198.5
1995	638	2.90%	614	6.5742	390.49	0.612	4.40%	1.45%	4	637	351.74	2.94%	29.04%	1211.21

1996	664	4.08%	635	23.9774	395.58	0.596	1.30%	-2.66%	8	663	373.04	6.06%	30.48%	1223.89
1999	664	0.00%	655	7.5846	355.76	0.536	-10.07%	-10.07%	1	664	394.49	5.75%	31.91%	1236.26
2000	653	-1.66%	643	7.4429	343.19	0.526	-3.53%	-1.91%	3	650	416.08	5.47%	33.35%	1247.61
2001	664	1.68%	650	14.0443	333.99	0.503	-2.68%	-4.29%	1	664	480.64	15.52%	37.66%	1276.27
2002	656	-1.20%	647	15.1689	343.99	0.524	2.99%	4.25%	2	656	502.12	4.47%	39.09%	1284.53
2003	655	-0.15%	648	10.3590	315.28	0.481	-8.35%	-8.21%	4	655	523.76	4.31%	40.53%	1292.27
2004	654	-0.15%	646	15.3146	314.44	0.481	-0.27%	-0.11%	6	654	542.83	3.64%	41.76%	1299.88
2005	658	0.61%	650	12.8599	322.74	0.490	2.64%	2.02%	2	657	562.12	3.55%	42.99%	1307.56
2006	660	0.30%	656	2.4054	316.45	0.479	-1.95%	-2.25%	1	660	582.88	3.69%	44.34%	1314.48
2007	654	-0.91%	653	0.9553	336.4	0.514	6.30%	7.28%	5	653	606.33	4.02%	45.89%	1321.29
2008	655	0.15%	654	9.0581	342.01	0.522	1.67%	1.51%	3	653	624.03	2.92%	46.99%	1328.02
2009	654	-0.15%	654	11.6994	322.65	0.493	-5.66%	-5.52%	3	654	645.12	3.38%	48.34%	1334.5
2010	656	0.31%	651	1.5231	319.25	0.487	-1.05%	-1.36%	8	655	669.78	3.82%	49.95%	1340.91
avg	410.36	12.76%	385.97	7.4454	228.768	0.517	18.73%	6.45%	5.971	414.514	304.04954	17.50%	25.33%	1006.374

CHAPTER 6
CITY SIZE AND LOCAL AIR QUALITY: EVIDENCE FROM CHINESE CITIES

6.1 INTRODUCTION

Rapid urban growth has been occurring the world over. The percentage of people lived in urban area is 30% in 1950, and until 2000 the fraction is 47%. By 2030 there will be 60% of the population live in cities according to this growth rate[54]. However, this rapid urban development, and the economic growth that fuels it, potentially comes at a cost. Urban environmental degradation is becoming an increasingly serious issue. The living quality of billions of people will be decided by the relationship between economic growth and the urban environment.

In this chapter we focus on China because this problem is more serious and obvious in China in recent years. China's urban economic growths substantially increase the income per capita and significantly raised the quality of life. Over the last thirty years, one quarter of the rural people who entered cities worldwide were in China. In order to vigorously promote China's urban development, a great number of building and housings constructed in a few months. Tens of millions of cars are registered. China faces a large amount of supply of electricity to accommodate the demand of growing high-income urban residents. In 2011, China's electricity consumption was about 4.5 trillion-kilowatt hours[55].

However, accompanying the on-going urban development are significant environmental consequences. The *Wall Street Journal* reports that the cost of environmental degradation in China was about $230 billion in 2010, about 3.5% of total GDP, representing a threefold

[54] United Nations, " World Population Prospects: The 2004 Revision Population Database" (esa.un.org/unpp [October 2005])

[55] http://www.bloomberg.com/news/2011-01-28/china-spower-demand-growth-may-slow-to-9-this-year-nea-says.

increase since 2004. The *Guardian* reports that "Chinese scientists have warned that the country's toxic air pollution is now so bad that it resembles a nuclear winter". A 2007 report entitled “The Cost of Pollution in China” concluded that studies have shown that outdoor air pollution leads to 350,000 to 400,000 prematurely die every year in China.[56] Two photos in Figure A6.1 and Figure A6.2 in Appendix also vividly show the severity of the smog in Chinese cities.

Most studies use local air pollution and greenhouse has production (GHG) to be two important indicators of urban air pollution. Nowadays, many Chinese cities suffer serious air pollution. In fact, WHO (world health organisation) points out twenty most polluted cities all over the world according to the urban concentration level of PM_{10}, among them, twelve cities locate in China (World Bank 2007b). In 341 Chinese cities with monitoring devices, 53% cities’ PM10 are higher than $100\ \mu g/m^3$; 21% cities reported PM_{10} exceed $150\ \mu g/m^3$. There are only 1% of cities’ air quality meet the criteria suggested by European Union’s standard that is 40 $\mu g/m^3$(World Bank 2007a).

In this chapter, we extend our analysis of city size and city growth from the previous chapters to the impact of city size on city environment. Specifically, we attempt to capture the relationship between city size and the local air pollution using 30 major cities in China (province capitals and 4 municipalities) for 10 years from 2003 to 2012. We measure city size by city population size, urban area size, city total GDP size, and city energy consumption size separately to investigate whether larger cities have better or worse air quality than smaller cities.

[56] The study was done by the World Bank in cooperation with the Chinese State Environmental Protection Administration.

The contribution of this chapter is threefold; firstly, unlike most of the previous literature focusing broadly on the economic growth and environment, we directly address the impact of city size on the local environment and use various indicators to represent city size. This partially fills the gap of the lack of studies on the impacts of city characteristics on the city environment. Secondly, we focus specifically on Chinese cities given the dramatic urbanisation process experienced in China in recent decades and given the lack of studies examining this issue for China. Finally, in light of the significant foreign direct investment into Chinese cities we also consider the environmental impact of industrial output by ownership category and differentiate between output that is domestically owned, foreign owned and HMT owned (Hong Kong, Macao and Taiwan).

The rest of this chapter is organised as follows. Section 2 thoroughly reviews the related literature on economic growth and environment, especially the city development and environment. Section 3 outlines the data statistics and the methodology employed. Section 4 discusses the estimation results for the three air pollutants (PM_{10}, SO_2, and NO2) and Section 5 concludes.

6.2 LITERATURE REVIEW

6.2.1 Measuring Urban Environmental Quality

The issue of how environmental quality is affected by urban growth has attracted a large number of studies from a range of different perspectives. However, there are several different definitions of the urban environment. Often we assume a green city to be one with clean air and water, tidy roads and parks, with resilience to natural disasters and with low incidence of infectious diseases. Green cities encourage 'green activities', for instance green cities will encourage people to use the relatively low-pollution public transportation. Researchers from different backgrounds have approached this issue from different angles.

Ecologist uses the concept of the Ecological Footprint to measure the urban environment. They emphasise the importance an urban ecological index and focus on the amount of people's consumption based CO_2 emissions. Wackernagel *et al.* (2002) compute the acreage needed to support human activity and to account for the waste from these activities: "in 1961 the human load accounts for 70% of the bearing capacity of earth's biosphere, however, in 1999 it accounts for 120%."

Public health experts find an alternative measure of the urban environment. They focus on people's health change due to urban air pollution, polluted water and other environment factors which may cause disease. According to this method, if within a city the morbidity of environmental related disease is low then we assume the city is a 'green city'. The specific method is to compute how much the residents' health will be improved if pollution is reduced.

From the economists' view, they tend to measure the urban environment by comparing the price of real estate across different cities at the same time point, or the price of real estate within a city over different time periods. If the environment is poor in a city, then the current residents tend to move out from the city and the potential residents will not consider moving into the city, which will result in a decline in the price of the real estate within that city. Similarly, cities with a higher quality of life will attract more migrants and then increase the price of real estate (and simultaneously lower the local wage). The price for real estate will adjust until there is no difference between living in a city with a good environment and a city with a poor environment. To maintain the equilibrium, the city with lower quality of life has to provide lower rent and better wage, i.e. 'differentiated compensation'. Examples of this approach are Bunten and Kahn (2014) who studies the impact of climate changes on urban real estate prices and Zheng, Kahn and Liu (2010) who examine house prices and air quality in China in 35 major cities.

However, in recent years, some economists explore the determinants of the air pollution due to the human economic activities directly, rather than examining real estate prices. Hiber and Palmer (2014) study 75 metro cities worldwide between 2005 and 2011 and attempt to find the determinants of air pollution for $NO_{2,}$ $SO_{2,}$ and $PM_{10.}$ Ebenstein *et al.* (2015) show the impact of growth and pollution on the life expectancy in China from 1991 to 2012.

6.2.2 Income Growth and the Urban Environment - Urban Environmental Kuznets Curves

6.2.2.1 Origins of the EKC and their Empirical Support

A large body of literature has examined the broad relationship between economic growth and the environment. The Environmental Kuznets Curve (EKC) captures the idea that economic development initially increases pollution levels but when income reaches a critical level the pollution level will decrease with further income growth. Over the last few decades, a significant body of work has attempted to study the complex relationship between income or economic growth and environment quality.

The EKC hypothesis emerged in the early 1990s at a time when environmentalists were concerned about the environmental consequences of the North American Free Trade Agreement (NAFTA). Firstly, from the scale effect perspective, the free trade would increase the economic activities in Mexico and therefore cause more pollution, other things being equal. Secondly, from the composition effect perspective, the stringent environmental regulations applied in the U.S. could result in dirty industries relocating to Mexico and degrading the Mexican environment. However, there is also a third mechanism, known as the technique effect. As Grossman and Krueger (1995) argue, free trade tends to attract more foreign direct investment and the employment of cleaner technologies. Also free trade would also increase Mexico's per capita income which may increase demand for environmental regulations. So, free trade and economic growth may result in changing techniques of production which could reduce environmental degradation, other things being equal.

Grossman and Krueger (1995) test the ambient sulphur dioxide and total suspended particulates using cross-country data. They estimate the reduced form as follows:

$$Air\ pollution_{it} = b_1 * GNP_{it} + b_2 * GNP_{it}{}^2 + X_{it} + e_{it}$$

Where GNP represents the gross national product of a nation and the term GNP^2 represents the nonlinear relationship between pollution and national per capita income, b_1 and b_2 are estimated coefficients using ordinary least squares. X_{it} is a vector of other covariates. If b_1 is positive and b_2 is negative then there will be an inverted-U shape between the pollution and income growth, i.e. an EKC. Grossman and Krueger calculate the turning point is located between $6,000 and $8,000 per capita, depending on the measure of the pollution.

Grossman and Krueger's (1995) study triggered a large volume of papers testing the EKC model. Selden and Song (1994) support the EKC and identify an inverted-U shaped relationship between per capita pollution and per capita GDP using cross-national panel data. Based on Grossman and Krueger (1995), they add another two air pollutants, oxides of nitrogen and carbon monoxide in addition to suspended particulate matter and sulphur dioxide. They confirm that per capita emissions of all four air quality indicators display and inverted-U shape with per capita GDP, although with substantially higher turning points.

Hilton and Levinson (1998) find evidence in support of the EKC using 48 countries over 20 years data for automotive lead emissions. They suggest that lead emissions show an inverted-U shape with respect to income, although the turning point level of income depends on the estimated function form and the time period used.

List and Gallet (1999) initially confirm the EKC hypothesis using state-level US data on sulphur dioxide and nitrogen oxide emissione for 1929 to 1994. They estimate the conventional reduced form of the relationship between per capita income and per capita emission for the U.S. states and find an inverted U-shape between income and pollution. They argue that the U.S. states are more homogenous than the sample of cross-countries in previous EKC studies.

McMillen (2004) supports the EKC by analysing the urban noise pollution. As noise pollution is often generated by airports, he tests the impact of the airport noise on the real estate values around the airport, using one of the world's busiest airports - Chicago O'Hare. When cities grow, there will be more cars, more manufacturing and more construction activities, all contributing to noise levels (Kahn, 2006). McMillen (2004) finds that real estate prices are about "9% lower within a 65 dB noise contour band of O'Hare in 1997". These represent the upward slope of the EKC, i.e. when cities grow the pollution will increase accordingly. However, when cities become richer they tend adopt more stringent noise regulations while richer urbanites also have self-protection methods, such as thicker windows etc. (Kahn, 2006). McMillen (2004) finds that "aircraft are becoming so much quieter that the airport can be expanded without causing a drop in local property values or tax bases". In addition, after a new runway was added to the Chicago O'Hare airport, the real estate price increased by $284.6 million in the densely populated area around the airport. These indicate that the noise level is controlled and decreases over time as cities become richer, reflecting the downward slope of the EKC.

Zheng and Kahn (2013) argue that there are two well-known theoretical models which may provide the micro foundations for the EKC hypothesis (Andreoni and Levinson 2001; Stokey 1998). Firstly, in Stokey's (1998) model, he assumes pollution "as a function of the aggregate economic activity, but it can be reduced by investing in the costly cleaner technology". When a country is poor, it tends to use the relatively low cost dirtiest energy and technology and thus pollution increases with economic growth. As income increases above a critical level, residents tend to be more concerned about the environment, thus the country tends to use more clean energy and cleaner technology. This mechanism sees the emergence of the EKC hypothesis.

Secondly, in Andreoni and Levinson's (2001) model they assume a representative consumer gets utility (U) from consumption (C) and loses utility from exposure to pollution (P):

$$U = U(C, P) = C - P$$

Pollution (P) can be expressd as two components, C and$-C^{\alpha}E^{\beta}$, as pollution will increase proportionally with respect to consumption, while it decreases with respect to pollution abatement $C^{\alpha}E^{\beta}$ which is related to consumption (C) and resources spent on environmental effort(E). The latter term is expressed using a standard Cobb-Douglas function $(0 < \alpha, \beta < 1)$:

$$P = P(C, E) = C - C^{\alpha}E^{\beta}$$

Andreoni and Levinson's (2001) also posit this representative agent spends resources/ income (M) on consumption(C), and emissions control(E), with the latter reducing pollution(P):

$$M = C + E$$

Andreoni and Levinson (2001) figure out the solution of “closed form solution of optimal pollution production” and find the function of income (M) to represent pollution. An Environment Kuznets Curve (EKC) appear if $\alpha + \beta > 1$.

Andreoni and Levinson’s (2001) and Stokey’s (1998) use a function of single sector production. They simplify the problem, and analyse the heterogeneous structure of industrial, capital vintage and the spatial distribution of activities in industrial and households.

6.2.2.2 Criticisms of the EKC

Several criticisms of the EKC hypothesis have been raised and focus on four major concerns about “irreversibility, short-run environmental degradation, the role of pollution havens, and the challenges posed by cross-border externalities” (Kahn, 2006). There is an implicit assumption in the EKC hypothesis that the pollution studied is reversible, it can be reduced or “undone” in the future. This is true for most of the pollutants, like the lead emissions studied in Hilton and Levinson (1998) or the noise studied in McMillen (2004). However, there does exist certain types of environmental damage that cannot be reversed, like the extinction of species.

The second concern is about the time it takes to reach the turning point for developing countries. For many developing countries, they are far below the turning point (lie far to the left the EKC inverted-U shape). In these countries rapid population growth and rising per capita income may result in substantial environmental damage and this damage will continue for several decades or even longer. Kahn (2006) points out that a turning point of $6,000 will imply a wait of 36 years before emissions begin to fall for a city with a per capita income of $3,000 and a growth rate of 2% per annum.

The third concern relates to the Pollution Haven Hypothesis (PHH) which questions the effectiveness of the EKC hypothesis. Some critics argue that the gains of the environmental improvement for developed countries to the right of the EKC turning point are partially provided by the environmental degradation in developing countries. As the developed countries make more stringent environmental regulation, a large body of dirty industry tend to move to the low-income countries with relatively lax environmental regulation. Therefore, developing countries may be trapped on the upward sloping portion of the EKC. The empirical evidence for the PHH is mixed. Some researchers explore international trade and FDI patterns and find support the hypothesis, especially for footloose industries, such as jewellery or office and computing machines (Ederington and Minier, 2003; Ederington, Levinson and Minier, 2005; Kahn, 2003b). However, other researchers find that the environment regulation in a country or region is not a significant determinant of industry relocation (Panayotou, 2000).

The final criticism of the EKC is the cross-border external effect which makes the EKC hypothesis less significant. For a problem such as acid rain, the pollution produced from one country may "externalized" onto neighbours. If a country's pollution drifts into neighbouring countries then the country of origin may have little incentive to control its pollution. Bradsher

(2002) shows that officials in Guangdong province, China have little incentive to limit or strictly regulate their rapidly growing manufacturing industry which has sharply increased particulate levels in Hong Kong.

Harbaugh *et al.* (2002) re-examine the EKC hypothesis and find little evidence for it in worldwide cities for the pollutants sulphur dioxide, TSP and smoke. They use the data used by Grossman and Krueger (1995) and combine with additional data. They find that the

EKC hypothesis is sensitive to moderate changes in both sample selection and econometric specifications (adding different countries, adding years, adding new control variables or using fixed rather than random effects). They not only show that the environment does not improve with economic growth beyond a certain point, but also argue that there is no evidence that the environment necessarily declines with economic growth. They conclude that for these pollutants there is no available empirical evidence to prove that economic growth is beneficial for the environment or harmful to it.

Stern and Common (2001) use a larger sample and a longer time series (1850 to 1990) and find different results for sulphur emissions in countries of different levels of development. Specifically, for developing countries (low-income countries), they find a monotonic function between income per capita and sulphur emissions, however, in the developed countries (high-income countries) they find an inverted-U shape between income and pollution i.e. the EKC hypothesis. They argue that the abatement of pollution tends to be time-related rather than income-related.

6.2.3 Population Growth and the Urban Environment

The EKC hypothesis focuses on the income growth of a city and its impact on the urban environment. However, other key urban characteristics may also affect the urban environment yet these are often neglected within the EKC literature, such as population growth, spatial growth and industrial structure etc. In this chapter we focus on the relationship between city size and air quality in Chinese cities, where by city size we mean city population size, spatial size, economic size and energy consumption size. This chapter attempts to capture the impact of some other key aspects of urban growth on environmental quality rather than the EKC's focus on income.

Kahn (2006) argues that urban population growth is a key driver of urban environmental degradation. In recent decades, cities attract a large body of people from rural areas seeking better jobs and a better quality of life. These inevitably increase the consumption of urban resources (real estate, transport, energy etc.) and produce more waste and pollution. This is especially true in developing countries where urbanisation tends to be more rapid.

6.2.3.1 Population and the Environment- Theoretical Studies

Theoretically, Zheng and Kahn (2013) summarise the emerging literature on cities and pollution to explore the causes and consequences of China's urban pollution. They summarise two urban system models which introduce spatial considerations into the EKC hypothesis. The classic EKC hypothesis is typically tested using national data and does not explicitly incorporate geography or deal with the city-level. Therefore, the two spatial urban models - the Rosen (1979) and Roback (1982) open system of cities model, and Alonso (1964), Muth (1969) and Mills (1967) (AMM) classic urban monocentric model - can link the EKC hypothesis to city level and explore the link between city growth and the environment.

Specifically, those spatial models help us to explore the migration of households or firms across cities or within a single city. In Rosen and Roback's model, the equilibrium is based on the adjustment of wages and rents according to the quality of life. If the quality of life is relatively lower in a city then the wages and rents tend to adjust to compensate for such disamenities (Albouy 2009).

The classic AMM model posits that all employment locates in the downtown business district. This urban environment model does not directly explore the industrial pollution, but use the proportional relationship between the industrial pollution and industrial employment and assume the pollution concentration is higher around the downtown areas. Thus, this offers the urbanites a trade-off of living in the centre downtown. They may have shorter commute times but have to expose to higher pollution levels. The AMM model predicts that if human agglomeration and jobs distribute more evenly throughout the urban area, there would be less "pollution hot spots" in downtown centre.

Zheng and Kahn (2013) adopt the ideas of non-spatial and spatial models together, and establish a framework attempting to find the determinants of urban environment degradation. They suggest that there might be two main sources of urban pollution, one is from the production activity, for instance industrial production process, and the other is from the residential activity, such as electricity consumption etc. They study the city and environmental problem from the supply and demand perspective, i.e. the supply of pollution production and the demand for greenness. The supply of pollution is from the production (industrial) and consumption based externalities as above; the demand for green cities is from the richer and better educated households The demand for green cities is higher for richer and higher educational level households, because they tend to pursue high quality of life and manage to reduce the exposure of environmental degradation risk. The choice of location is a kind of behavior of self-protection. If households are different in the sensitivity of exposure to pollution, then the most sensitive families would choose to live in relatively low-pollution level cities or cleaner locations in some specific part of the city (Ehrlich and Becker, 1972). They also introduce the impact of government on the mitigation of urban pollution.

6.2.3.2 Population and Environment- Empirical Studies

Kahn (1999) studies the U.S. Rust Belt's decline in the 1960s and 1970s and finds that this industrial decline significantly improved environmental quality in heavily industrial cities such as Pittsburgh, Pennsylvania and Gary, Indiana.

Van Der Waals (2000) tests whether the form of compact city can reduce or prevent urban environmental degradation. The pollutants they test for are CO_2 and NO_2 caused by mobility; noise, odour and local air pollution; fragmentation of natural areas. They find that in the short run, using the compact city policy to solve the environmental problem is limited. However, a policy of "concentrating urbanisation" might be helpful.

Kahn (2006) examines the impact of population growth on three of the leading sources of urban environmental degradation: air pollution, water pollution, and solid waste. He shows a positive correlation between levels of particulates or sulphur dioxide and city size (population). He use U.S. air pollution data from 1973 to 2000 and estimates that controlling for the monitoring station fixed effects (which absorbs time-invariant influences such as the difference in geography, average wind speeds or climate conditions) if a country's population increases by 10 percent, and then the ambient particulate levels would increase by 4.4 percent accordingly. And if the country's population is held constant, the time trend shows that ambient particulates decline by 2.6 percent each year, which is probably due to the strengthening of environmental regulation and/or the use of cleaner technology. With respect to water pollution, as city populations increase the demand for basic services also increases. However, Kahn (2006) argues that especially in developing countries, most of the migrants are from rural areas and are too poor to afford the city services. This makes the existing service over-taxed and its quality falls. Finally, for solid waste, with growing population in a city, the amount of domestic refuse would increase sharply. It becomes even worse when the income growth is not significant, because there is no incentive or effective way to deal with the waste. Kahn argues that this is most severe in cities where property rights are poorly defined.

Dasgupta *et al.* (2004) show that in addition to population size, a city's climate conditions, governance, and income all determine the local air quality in a significant manner.

Kahn and Schwartz (2008) argue that a city's pollution level at a point in time is a function of scale, composition and technique effects. Specifically, the scale effect is the sheer count of people or jobs or aggregate city level economic activities. The composition effect refers to the industrial composition and capital stock which often significantly affects the pollution level as heavy industries are typically more pollution intensive than service industries. Finally, the technique effect refers to the nature of the production process which affects emissions per unit of economic activity. Kahn and Schwartz (2008) suggest that composition and technique effects can offset the externality costs of population growth.

Kahn (2010) studies the relationship between urban congestion/ pollution/ crime and city size in the U.S. He examines the population elasticity of producing local public disamenities such as crime, pollution, and commuting times. In terms of pollution, he studies CO_2, PM_{10}, TSP, and NO_2 using micro data series (1980, 1990, and 2010) of individual level and household level data from the U.S. census. The approach of his study is to estimate the effect of city attributes (commute times, urban air pollution, and crime) on the city population size. He finds that firstly, big cities have experienced sharp reductions in ambient pollution in recently years; secondly, the "big-city premium" is more obvious in north-eastern cities in terms of pollution and crime.

More recently, Kahn and Walsh (2014) investigate the city and environment from the urban environmental amenity dynamics point of view. They suggest that a city's demographics, industrial structure, and the environment policies are the determinants of a city's production of pollution. They emphasise the role of local public good dynamics on the spatial equilibrium cross-city. More specifically, they focus on the new local and federal policies such as "local zoning and differential enforcement of the Clean Air Act".

Hiber and Palmer (2014) study the determinants of air pollution concentration for NO_2, SO_2, and PM_{10} for 75 metro cities worldwide between 2005 and 2011. They also explore how the changes of urban form and transportation mode affect pollution concentrations. They find that, within cities, the increasing car density leads to a reduction of air pollution concentration for NO_2 and SO_2 (results are largely found in non-OECD countries), which might be partially because the increasing car use decentralises the residential and economic activity. They also find a complex relationship between income and pollution concentrations. Overall, their results indicates that "densely populated polycentric cities may be greener and healthier than comparable monocentric ones", decentralization of the population or car density is beneficial to air pollution concentrations.

In conclusion, population growth indeed affects the environment because of the increasing consumption of resources and the construction of cities. The relationship between city population growth and the urban environment is important for urban planners who need to tackle the environmental degradation which may accompany rapid urban population growth.

6.2.4 China's urban environment

Over the past two decades China has experienced an economic growth rate of approximately 8% each year and has seen average income increase sharply from $894 per capita to $9,087 per capita[57]. However, this rapid economic growth may come at a cost as environmental degradation is often an unintended by-product of economic growth. For developing countries such as China, it is essential for government or urban planners to understand the relationship between urban development and environment when formulating urban policies.

[57] China Statistical Yearbook 2013.

In recent years China's smog has increasingly attracted the world's attention. The *Wall Street Journal* reports that the cost of environmental degradation in China was approximately $230 billion in 2010, or 3.5% of the gross domestic product. This number is three times the amount environmental cost estimated for 2004. The *Guardian* reports that "Chinese scientists have warned that the country's toxic air pollution is now so bad that it resembles a nuclear winter". Many cities in China have high levels of air pollution. For the ambient particulate concentration $PM_{10,}$ as mentioned above twelve out of twenty most polluted cities are in China around the world (Word Bank 2007b).

Zheng and Kahn (2013) survey the literature on the relationship between China's economic growth and its urban environment consequences. They analyse the problem from the supply and demand perspective. For the supply of pollution, they suggest that there are two key issues; the first is the urban pollution source. They summarise two basic categories of pollution sources, the production based pollution (for instance, industrial production) and the consumption based pollution (for instance, residential electricity and vehicle driving); the second issue relates to the different categories of pollution indicators, local (air pollution) and global (GHG, greenhouse gas) externalities. In this chapter, we mainly focus on the local pollution or air pollution in this chapter, with respect to the GHG emissions we summarise some of the literature in the end of this section.

The demand for ‘greenness’ stems from the fact that as people getting rich and being well educated they prefer more information and products to reduce their risk of exposure to pollutions. Therefore, the location choice model is often discussed in the literature about the urban environment. The choice of location is a kind of self-protection behavior. If the difference in household is their sensitivity of pollution, then the most sensitive families tend to choose self-protection behavior of moving to the cities with low-pollution level or certain clean location of the city (Ehrlich and Becker, 1972).

In addition, Zheng and Kahn (2013) also discuss the role of government in abatement of the environment degradation. They suggest that air and water are public goods with no market price or personal cost of pollution, which leads to a Coasian bargaining solution. Selden and Song (1995) argue that the J-curve theory predicts that the increasing of per capita income are able to push the stringency of the environmental regulations. Hilton and Levinson (1998) also points out that a country tend to be more strict with the allowable pollution production when it getting richer.

In conclusion, the above literature considers China's urban pollution to be determined by city population growth, industrial activity, driving and urban form, power generation, winter heating, and global externalities such as natural resource extraction and GHG emissions.

Nonetheless, the extent to which that the economic growth will affect the local environment depends on the scale, composition, and technique effects. In addition, the geography of the pollution activities matters as well. Thus in the following section, we analyse the scale, composition, and technique effects respectively.

6.2.4.1 The Scale Effect- Urban Population Growth

The scale effect refers to the impact of sheer amount of urbanites or jobs or aggregate city level economic activities on the urban environment. In China, as a developing country, there are a large number of rural residents seeking job opportunities and better quality of life in cities (Harris and Todaro, 1970). Fan (2005a, 2005b) have addressed the large rural-urban migration trends. The urbanisation rate has been raised from 26% in 1990 to 51% in 2010. Census 2010 shows that 78% of the migrants in cities are from rural areas. By 2030, it is predicted that 70%-80% of Chinese residents will live in urban area according to this rate of urban growth.

Zheng and Kahn (2013) argue that there might be two effects of people moving into cities, which can offset each other. On the one hand, the increasing migrants within a fixed geographic area will raise the population density, which might forces people living in smaller houses and thus reduce the energy consumption. On the other hand, people moving into cities tend to be richer than before, which might increase the energy consumption. Future more, Zheng *et al.* (2011) address these two effects and find that the latter (income effect) may have more influence. This means that the net effect of the migration on environment is negative.

Studies on U.S. find that urbanisation has positive environmental effect on rural areas where the density is lower when people moving into cities. Pfaff (1999) points out the positive environmental effect to the forest of Northeast when urbanisation on process in the 19th century. However, in China, Ebenstein *et al*. (2011) find a rise of local nitrogen levels and subsequent increasing water pollution due to the large amount of migrants moving into cities.

As this reduces the labour supply in rural area and increases the usage of chemical fertilizer in the production process.

It is also worthy to mention the role of Chinese "*hukou"* system, which significantly restricts rural-urban migration for certain periods from 1958 to 1980s, thus may restrict the scale effect during this period. [58] Because the object of our analysis is China, and studies are often interested in whether its unique policies may affect the results. However, the period used in this chapter is the relaxation period of population migration (2003-2012), thus this unique policy in China may not affect our results.

6.2.4.2 The Composition Effect- Urban Industrial Production

The composition effect refers to the industrial composition and capital stock which often significantly affects the pollution level as heavy industries are typically more pollution intensive than service industries.

Ebenstein (2012) studies the industrial activity and the environment and confirms that industrial activity has led to a severe deterioration in water quality in China's lakes and rivers, using 145 Disease Surveillance Points (DSP) in China and water quality measures from China’s nationwide monitoring system. A deterioration of water quality by a single grade (on a six-grade scale) increases the digestive cancer death rate by 9.7%.

[58] Basically the "hukou" as a registration system denoting people's original hometown by urban or rural status was launched in 1950s. After the "Economic Reform" from 1980s this regulation has been relaxed substantially in order to accommodate the transition to a market economy. In 1990s, the population mobility has been pushed up as the rising of the housing and labour market in urban area.

Cole and Elliott (2003) explore the relationship between trade and environment, and examine the determinants of the trade-induced composition effect, using SO_2, NOx, CO_2 and BOD (biochemical oxygen demand) as the dependent variable separately (SO_2 and NO_2 observations at 5-year intervals for 26 countries from 1975-1990; CO_2 and BOD observations are yearly for 32 countries from 1975-1995). Specifically, the literature on trade and the environment often divides the impact of trade liberalisation on the environment into scale, composition and technique effects. Cole and Elliott (2003) examine whether the composition effect changes because of trade liberalisation (originally due to the differences in capital-labour endowments or differences in environment regulations). They find that the composition effect due to trade is in fact not obvious, and different pollutants tend to have different sign and magnitude of these effects. Such as for SO_2 emissions, if income increase by 1% due to trade, and then this would lead to 1.7% per capita emission reduction. But with the positive trade intensity elasticity, the final net outcome is ambiguous; trade liberalisation tends to reduce BOD emissions and increase NO_X and CO_2 emissions.

6.2.4.3 The Technique Effect- FDI and Industrial Output from Foreign Firms

The technique effect refers to the effect of changing techniques of production on emissions per unit of economic activity. Many studies examine the impact of multinational firms on environment and assessed whether the multinational firms producing less pollution comparing to domestic firms. For instance, Eskeland and Harrison (2003) and Cole, Elliott, and Strobl (2008) empirically suggest that "multinational firms in developing countries are less pollution intensive than their domestic counterparts ". While some other studies find no such evidence to support this argument (Pargal and Wheeler, 1996).

The possible explanations that multinational firms might be less pollution producing than domestic firms in developing countries are from three perspectives. Firstly, multinational firms tend to employ more advanced technologies; secondly, they might have cleaner production process; thirdly, in terms of management, they may have more developed and comprehensive management system. In addition, multinational firms tend to have large proportion of export to OECD countries where the environmental regulation is more stringent.

The pollution haven hypothesis suggests that less stringent environmental regulations in developing countries may attract more FDI and thus more pollution. Nevertheless, the evidence supporting this hypothesis is few (Xing and Kolstad, 2002; Eskeland and Harrison, 2003; Cole and Elliott, 2005). Previous literatures show little clear understanding of the effects of FDI on local air quality. Neither for the city-level studies on environment, i.e. how city-level characteristics affect local air quality. Thus, this chapter attempts to show the impact of one of the city characteristics- city size- on the local air quality, and to what extend that different ownerships of firm can affect the local air quality, using data for 30 cities between 2003 and 2012. Because our data are available to be separated different type of firm's output into Hong Kong, Macao and Taiwan origin firms.

Empirically, for evidence in China, Zheng *et al.* (2010) find cities experiencing inflows of FDI have lower air pollution levels than observationally identical cities using a panel data set (35 Chinese cites from 1997 to 2006).

6.3 DATA AND METHODOLOGY

6.3.1 Data description

In this chapter we use pollution data from the *Environmental Statistics Yearbook (2003-2012)* published by China's Ministry of Environmental Protection (MEP), we then combine this with city demography data and industrial information from the *Urban Statistical Yearbooks (2003-2012).* Although the *Urban Statistical Yearbooks* report demography and industrial data for around 600 cities (including county-level cities), the *Environmental Statistics Yearbooks* provide urban pollution data (PM_{10}, and concentrations of SO_2, NO_2) for 30 cities. Therefore, we use a panel of data for 30 major Chinese cities from 2003 to 2012. Table 6.1 below shows the units and some statistics for the variables we use in this chapter, and Figure 6.1 below illustrates the location of these 30 major cities (province capitals) in China.

Table 6.1 Summary Statistics

Variable name	Definition	Year	no. of cities	Obs	Mean	Std.Dev	min	max
PM10	Particular matters (milligram/cu．m，day)	2003-2012	30	300	0.10	0.03	0.03	0.19
SO2	Sulphur dioxide (milligram/cu．m，day)	2003-2012	30	300	0.05	0.02	0.01	0.15
NO2	Nitrogen dioxide (milligram/cu．m，day)	2003-2012	30	300	0.04	0.01	0.01	0.07
AvgTemp	annual average temperature (℃)	2003-2012	30	300	14.48	5.11	4.50	25.40
MaxTemp	annual maximum temperature (℃)	2003-2012	30	300	36.82	2.82	28.90	43.00
MinTemp	annual minimum temperature (℃)	2003-2012	30	300	-9.87	10.77	-32.50	36.00
AvgHumi	Annual average humidity (%)	2003-2012	30	300	65.16	9.73	42.00	85.00
AvgSun(hour)	annual average sunshine (hour)	2003-2012	30	300	1,976.36	504.05	681.60	3,093.30
AvgPre(millimeter)	annual average precipitation (millimetre)	2003-2012	30	300	920.74	701.81	74.90	9,364.00
pop	population (1000 persons)	2003-2012	30	300	4,259.76	3,462.66	718.20	17,791.00
popden	population density (person/ sq km)	2003-2012	30	300	1,649.82	1,339.25	223.31	11,449.30
built-up	built-up area (sq km, from total urban area)	2003-2012	30	300	351.93	269.86	43.00	1,350.00
urbanroad	end of year urban road area (1000 sq m)	2003-2012	30	300	42,130.03	33,769.00	4,900.00	214,900.00
GDP	total GDP for a city within a year (million RMB)	2003-2012	30	300	26.91	33.41	0.95	199.45
GDPpercap	GDP per capita (RMB)	2003-2012	30	300	47,480.62	22,311.12	9,667.00	121,553.00
GDPgrow	GDP annual growth rate (%)	2003-2012	30	300	13.61	2.68	5.80	26.60
secondary/GDP	Annual average percentage of secondary industrial (%)	2003-2012	30	300	42.89	8.10	19.61	60.49
service/GDP	Annual average percentage of service sector (%)	2003-2012	30	300	54.77	8.03	37.95	78.66

AvgWage	average annual wage per employee (RMB)	2003-2012	30	300	638,601.35	1,737,525.91	16,911.09	14,934,563.00
inoutput	total industrial output (1000RMB,current year price)	2003-2012	30	300	354,000,000.00	509,000,000.00	5,037,560.00	3,200,000,000.00
dooutput	industrial output by domestic firms (1000RMB,current year price)	2003-2012	30	300	208,873,091.51	234,337,196.65	4,843,140.00	1,219,969,250.00
HMToutput	industrial output by HMT firms (1000RMB,current year price)	2003-2012	30	300	37,127,691.18	78,157,486.88	23,040.00	544,918,800.00
Foroutput	industrial output by foreign firms (1000RMB,current year price)	2003-2012	30	300	110,580,219.01	224,227,495.68	171,380.00	1,446,633,460.00
elec	total electricity usage (1000kwh)	2003-2012	30	300	2,103,960.37	2,308,064.72	152,027.00	13,534,500.00
inelec	industrial electricity usage (1000kwh, from the total)	2003-2012	30	300	1,259,142.93	1,410,999.85	41,658.00	8,057,600.00
resielec	residential electricity usage for living (1000kwh, from the total)	2003-2012	30	300	321,606.36	327,217.96	18,790.65	1,873,800.00
coalgas	incl. artificial gas and natural gas (10000 cbm)	2003-2012	30	300	137,473.73	283,193.09	618.00	1,976,114.00
coalgasresi	coal gas used for household (10000 cbm)	2003-2012	30	300	22,480.39	32,345.66	147.00	205,347.00
LPG	liquefied petroleum gas (LPG, ton)	2003-2012	30	300	110,813.74	163,513.62	2,110.00	1,087,766.00
LPGresi	LPG used for household (ton)	2003-2012	30	300	58,878.81	75,052.87	360.00	452,138.00

Notes:

All data use 'city proper'-'downtown' in the statistical yearbook

Urban area: refers to the total urban administrative area including the water surface area. The calculation is based on the administrative criterion.

Built-up area: refers to, in the urban area, the government approved land use and the real developed non-agriculture industry land use, including the concentrated contiguous land use in the downtown and the land area highly connected with the city development scattered in the surrounding suburban areas which have the basic urban public facilities (e.g. airport, sewage disposal plant, communication station etc.).

Industrial output: refers to the total value of the industrial product within a year, including 1) the value of the end-products 2) the income from the external processing 3) the value gap between the beginning of the year and the end of the year for the self-made semi-finished product.

Total electricity usage:includes all type of electricity use, e.g. rural area, industrial, transportation, urban residential usage. In all the types, the electricity usage includes the amount that Electricity Company sold, the private electricity plant usage by itself, and the amount that private electricity plant sold to the adjacent customer.

Fixed assets: comprises 1) Urban fixed assets investment, including the investment at 500 thousands RMB and above by any type of firms, administrative unit and private activity. 2) Rural fixed assets investment, including within the rural area any fixed investment by no-agriculture people.

Real estate investment: refers to the real estate development and operation activity by any type of registered real estate firm, including building, demolition and repair of residential houses, plant, storage, restaurant, hotel, holiday village, office building etc. and related services; land development projects including road, water supply, drainage, electricity supply, heating supply, communication, land formation; not including the pure land transaction.

Home investment: refers to the buildings for dwelling only, including the villa Dom, apartments, staff and students accommodation etc. not including the area for people's air defence, or basement.

Urban road area: refers to the road paved properly, the width of the road is above 3.5m (incl. 3.5m), including high-class, sub-high-class and normal road, not including the 'Hutong' (lane) which width smaller than 3.5m.

Figure 6.1: 30 Major Cities (Province Capitals) in the Sample.

Source: China administrative. svg (public domain).
Note: we exclude Tibet and Hong Kong, Macau and Taiwan in our sample because of insufficient data.

6.3.1.1 Pollution Concentrations versus Emissions

In this chapter we focus on pollution *concentrations* data, as opposed to emissions, as we are analysing the exposure to pollution within a city.

Cole and Elliott (2003) suggest that the different measurements of the same pollutant are important in testing for the pollution related issues. The literature on the Environmental Kuznets Curve shows that results may differ depending on whether the pollutant is measured in terms of concentrations or emissions. For instance, Selden and Song (1994) find that the estimated EKC turning point for city-level concentrations is lower than that for national emissions, which might be because pollution concentrations are easier to reduce or disperse than the pollution emissions.

More specifically, Cole and Elliott (2003) argue that concentrations and emissions of pollution provide different information. City-level pollution concentrations can provide more information about the local impact of a particulate pollutant on human health or quality of life, because residents are exposed to the pollutant concentration within a city. In contrast, the national or city emissions provide an aggregate amount of a pollutant emitted within a certain period. This may be a poor measure of pollution exposure as wind, precipitation, the height of chimneys and many other factors can limit the extent to which emissions increase pollution concentrations in a particular area. Therefore, in this chapter, we use air pollutant concentration data within each city rather than pollutant emissions, which show the urbanites' direct exposure to the air pollution.

In addition, there are also some natural problems of concentration data because of the heterogeneity of the monitoring sites - the site specific effects (Cole and Elliott, 2003). Cole and Elliott (2003) point out that it is necessary to control for the nature of the monitoring (observation) sites, for instance the location of the sites (city-centre, suburban, or rural); the type of measuring equipment; the ambient climate condition of the sites (temperature, rainfall etc.). It is often required to add in some dummy variables to capture the site-specific effects. With respect to this problem, since the 30 cities in our data are province capitals, they are the focus of the environmental monitoring activity by central government and tend to have similar levels of pollution control and air quality monitoring equipment (all of the monitoring equipment is controlled and monitored by China's Ministry of the Environmental Protection (MEP)). We include the climate condition variables (maximum and minimum temperature, average precipitation, average humidity, and average sunshine hours) and employ the city-fixed effects estimation to control for the site specific effects.

6.3.1.2 Data statistics

Table 6.1 shows the mean and variance for each variable among all of the observations and the detailed explanation for the variables. Table 6.2 and Table 6.3 shows the snapshots of the data we used in this chapter for one year 2012. Table 6.2 shows the pollution statistics and Table 6.3 shows the climate statistics, all in descending order of population.

Table 6.2 Pollution Statistics for 30 Major Cities in 2012

Chinese name	City	Region	PM10	SO2	NO2	pop	popden	GDP	GDPpercap	built-up area
重庆	Chongqing	West	0.09	0.037	0.035	17791	601.26	87.6008	49486	1052
上海	Shanghai	East	0.071	0.023	0.046	13584	2635.09	199.4537	86995	886
北京	Beijing	East	0.109	0.029	0.052	12265	1006.4	176.17	89659	1261
天津	Tianjin	East	0.105	0.048	0.042	8125	1098.12	119.0678	93173	711
广州	Guangzhou	East	0.069	0.022	0.049	6780	1764.17	124.5493	111704	1010
郑州	Zhengzhou	Mid-land	0.105	0.051	0.046	5872	5813.86	28.11784	63850	355
西安	Xi'an	West	0.118	0.04	0.042	5728	1598.99	36.5947	56203	375
成都	Chengdu	West	0.119	0.033	0.051	5542	2551.57	57.31731	74167	516
南京	Nanjing	East	0.102	0.033	0.051	5533	1169.11	64.6692	88525	637
沈阳	Shenyang	North-East	0.092	0.058	0.036	5221	1504.24	52.66876	82878	430
武汉	Wuhan	Mid-land	0.097	0.03	0.054	5130	1887.38	64.8495	94474	506
哈尔滨	Harbin	North-East	0.094	0.036	0.047	4714	665.2	30.05325	63747	367
杭州	Hangzhou	East	0.087	0.035	0.053	4454	1451.86	62.13249	98697	453
长春	Changchun	North-East	0.087	0.03	0.044	3630	757.95	31.29274	90987	434
济南	Jinan	East	0.104	0.055	0.041	3522	1081.27	36.32829	82369	363
长沙	Changsha	Mid-land	0.088	0.028	0.044	2979	1559.69	40.33603	110102	316
太原	Taiyuan	Mid-land	0.08	0.056	0.026	2841	1926.31	21.12533	60982	300
南宁	Nanning	West	0.069	0.019	0.033	2746	417.95	17.8717	65301	242
昆明	Kunming	West	0.067	0.034	0.036	2727	590.79	23.07844	59163	298
乌鲁木齐	Urumqi	West	0.145	0.058	0.068	2518	262.96	19.88324	60924	384
石家庄	Shijiazhuang	East	0.098	0.058	0.04	2471	8074.84	15.73539	53381	210
南昌	Nanchang	Mid-land	0.088	0.045	0.039	2252	3649.27	20.65285	73711	208
贵阳	Guiyang	West	0.073	0.031	0.028	2246	934.5	13.0549	41776	230
合肥	Hefei	Mid-land	0.098	0.019	0.027	2222	2401.84	27.47578	75795	378
兰州	Lanzhou	West	0.136	0.041	0.039	2064	1264.58	13.43543	47240	199
福州	Fuzhou	East	0.06	0.008	0.035	1921	1075.36	20.83591	70060	240
海口	Haikou	East	0.034	0.006	0.019	1616	701.04	8.18755	38634	124
呼和浩特	Hohhot	Mid-land	0.091	0.051	0.037	1220	590.85	17.34211	85102	174
银川	Yinchuan	West	0.099	0.044	0.037	1002	433.71	7.135115	54053	135
西宁	Xining	West	0.105	0.035	0.026	918	1799.22	5.928942	52852	75

Source: Compiled from Urban Statistic Yearbook 2013
To be noted that we use ‘city proper’ (within urban area, not including the subordinate suburbs) data in this thesis, not the whole city area.
For the unit for each variable, see Table 6.1.

PM_{10}, NO_2, SO_2 are particular matters, sulphur dioxide, and nitrogen dioxide respectively, measured as milligram/cu．m，day.

Table 6.3 Climate Statistics for 30 Major Cities in 2012

Chinese name	City	Region	AvgTemp	MaxTemp	MinTemp	AvgHumi	AvgSun	AvgPre	pop
重庆	Chongqing	West	18.3	39.9	1	72	812	1104.4	17791
上海	Shanghai	East	17.1	38	-2.5	68	1420.4	1435.8	13584
北京	Beijing	East	12.9	38	-13.7	51	2450.2	733.2	12265
天津	Tianjin	East	12.5	37.6	-13.3	57	2174.4	755.3	8125
广州	Guangzhou	East	21.7	36.8	2.5	82	1471.2	1813.9	6780
郑州	Zhengzhou	Mid-land	15.5	41	-7.6	53	1883.7	498.7	5872
西安	Xi'an	West	14.6	39.9	-9.3	62	1546.6	426.7	5728
成都	Chengdu	West	15.9	35.5	-3.4	78	780	610.9	5542
南京	Nanjing	East	16	37.2	-7.2	68	1939.1	917.2	5533
沈阳	Shenyang	North-East	7.4	32.4	-28.8	69	2577.2	786	5221
武汉	Wuhan	Mid-land	16.4	37.5	-6.5	81	1553.9	1415.5	5130
哈尔滨	Harbin	North-East	4.6	33.2	-31.3	67	1773.8	740.8	4714
杭州	Hangzhou	East	17.1	38.5	-3.3	71	1520.5	1728.8	4454
长春	Changchun	North-East	5.2	31.5	-27.6	63	2438.3	718.3	3630
济南	Jinan	East	14.3	37.5	-11.8	55	2146	569.1	3522
长沙	Changsha	Mid-land	17.6	37.7	-2.4	76	1493.6	1730	2979
太原	Taiyuan	Mid-land	10.7	36.4	-16.1	51	2618.6	427.8	2841
南宁	Nanning	West	21.4	37.1	3.1	80	1295.7	1086.6	2746
昆明	Kunming	West	16.3	30.9	-2.3	67	2554.2	802.1	2727
乌鲁木齐	Urumqi	West	7.4	36.2	-29	53	2864.6	286.9	2518
石家庄	Shijiazhuang	East	14	39.3	-10.9	55	2288.2	649.4	2471
南昌	Nanchang	Mid-land	18	36.9	-2.9	77	1622	2059.8	2252
贵阳	Guiyang	West	13.7	31.9	-4.1	85	681.6	1226.4	2246
合肥	Hefei	Mid-land	16.5	37.8	-6.6	73	1912.8	936.4	2222
兰州	Lanzhou	West	10.5	35.3	-17.1	53	2322.8	293.9	2064
福州	Fuzhou	East	20.2	39.2	0.8	75	1291.3	1913.4	1921
海口	Haikou	East	24.6	36.4	9.2	82	1766.4	2094.3	1616
呼和浩特	Hohhot	Mid-land	7.2	34.7	-23.9	47	2677.8	551.4	1220
银川	Yinchuan	West	9.8	34.8	-20.1	48	2728.3	292.7	1002
西宁	Xining	West	5.2	28.9	-23.8	59	2655.2	446.1	918

Source: Compiled from Environmental Statistic Yearbook 2013

PM_{10}, NO_2, SO_2 are particular matters, sulphur dioxide, and nitrogen dioxide respectively, measured as milligram/cu．m，day.

Particulate Matter (PM_{10})

Figure 6.2 A-D below shows the evolution of PM_{10} concentration for cities in different regions in China for the period 2003-2015. Particulate Matter (PM) refers to particles suspended in the air, consisting mainly of sulfate, nitrates, ammonia, sodium chloride, black carbon, mineral dust and water. It is more dangerous than other pollutant in affecting people's health, especially those with a diameter of 10 microns or less ($\leq PM_{10}$), which can be inhaled into deep lungs and cause cardiovascular and respiratory diseases, as well as cancer. The Environmental Protection Agency (EPA) finds that it comes from a variety of sources, such as diesel trucks, woodstoves, power plants, etc. and it can be generated directly "when gaseous pollutants such as SO_2 and NOx react to form fine particles".

The data we used is the annual mean concentration of PM_{10} for each city from 2003 to 2012. The WHO's (World Health Organization) Air Quality Guideline for PM_{10} is 20 $\mu g/m^3$ annual mean, or 50 $\mu g/m^3$ 24-hour mean. MEP (Ministry of Environmental Protection) for China reports the PM_{10} in milligram/ m^3 which is 1000 times of μg (micro gram). From Figure 6.2 we can observe that all of the major 30 cities in China exceed more than twice the WHO guidelines for air quality (Chinese cities' annual mean PM_{10} concentration are all above 50 $\mu g/m^3$ except for Haikou, Hainan province). Regionally, the North-East regional cities are the most polluted, with PM_{10} concentrations over 80 $\mu g/m^3$, four times the WHO guideline, which might be because the North-East region is traditionally the centre of heavy industry.

The time trend for most of the cities are decreasing and generally the decreasing trend in Eastern cities is more obvious than that in Western cities, which may reflect greater affluence in the East and a greater demand for environmental regulation. The most significant reduction of PM_{10} occurred in the cities in the North-East (3 province capitals), but the PM_{10} values are still around 90 $\mu g/m^3$, more than four times the WHO guideline.

Over this period while the economic growth did not slow down (average GDP annual growth rate still around 7%-8%, China Statistical Yearbook), but the pollution of PM_{10} generally decreases over these cities. This might due to the explicit environment policies enforced in these major cities (province capitals) from around 2005 until now, including optimize the structure of energy consumption and the distribution of industries in city adjacent areas; promote the consumption of natural gas and similar clean energy; develop public transportation, etc.

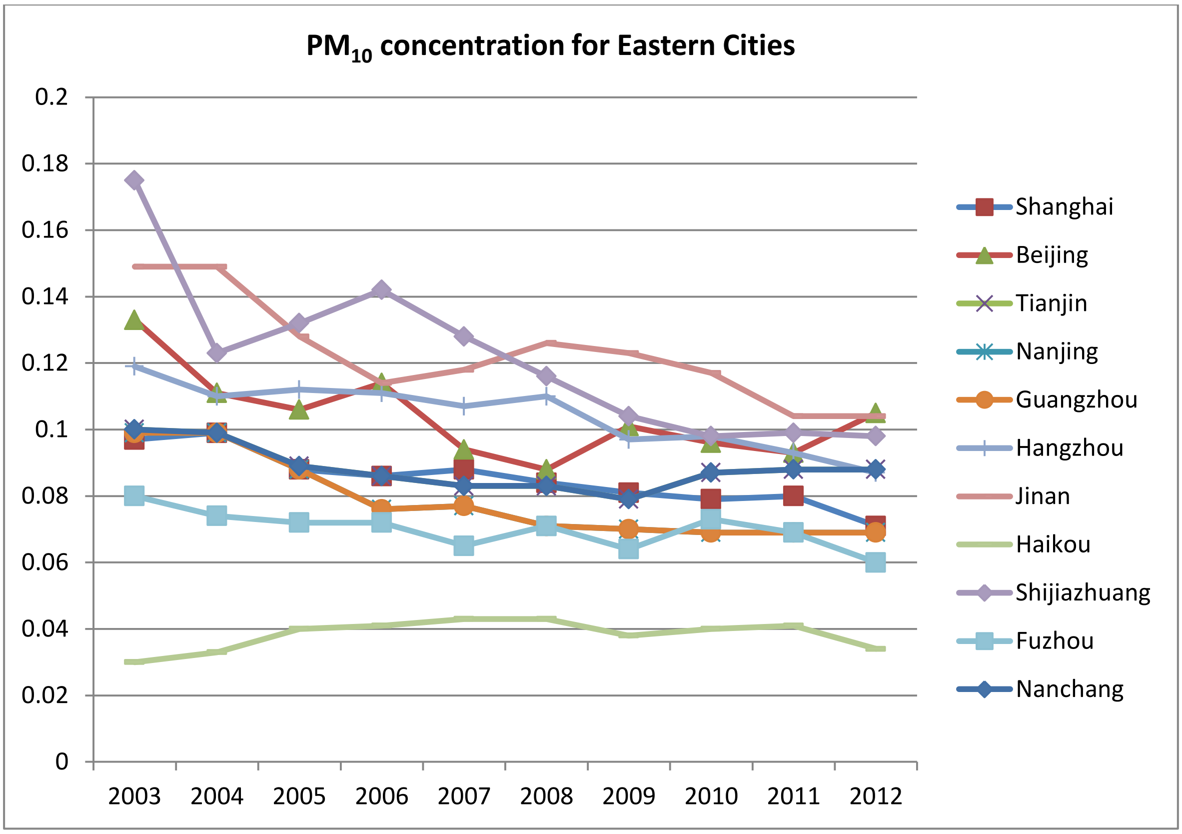


Figure 6.2- A

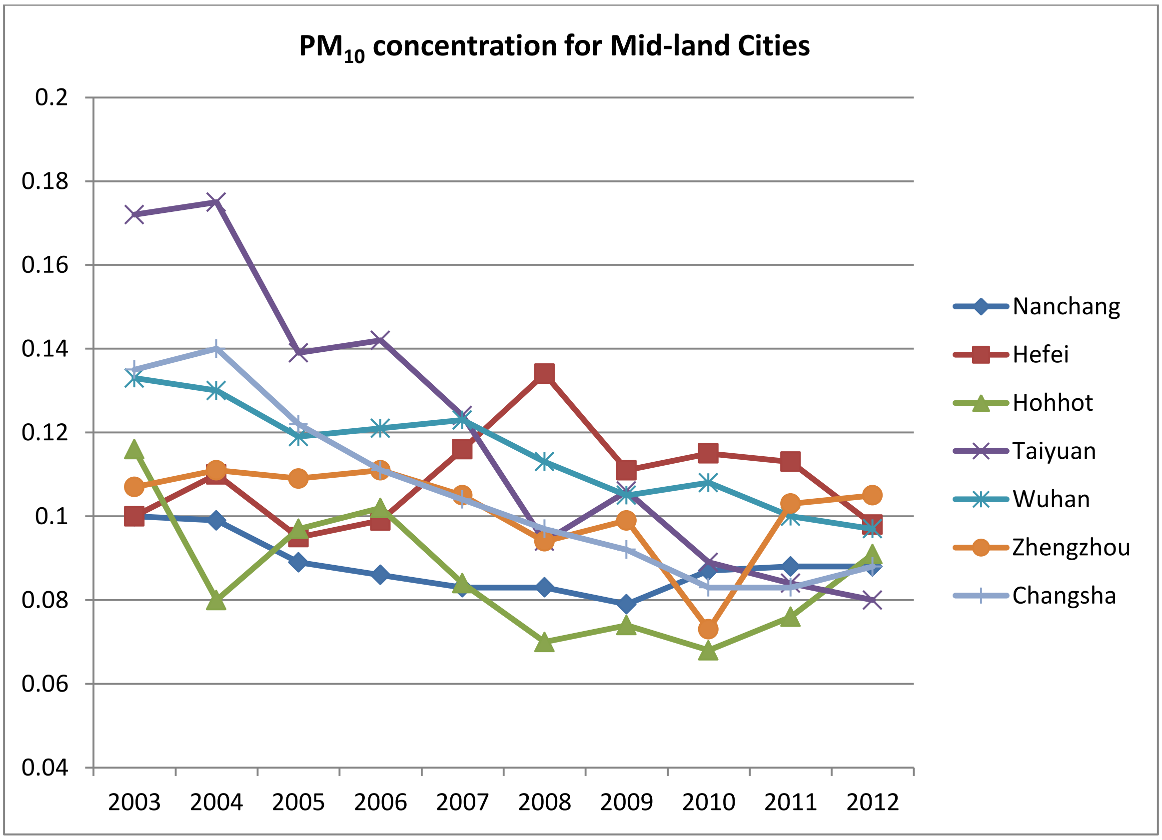


Figure 6.2- B

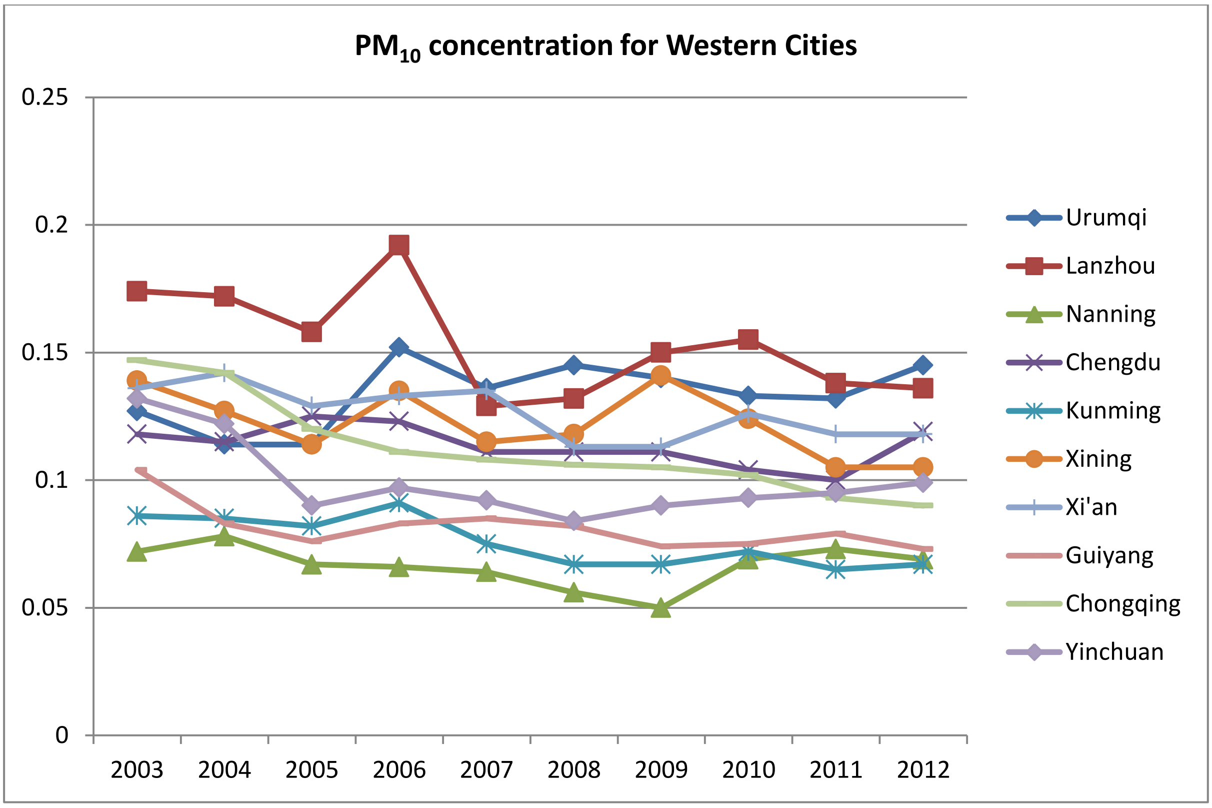


Figure 6.2- C

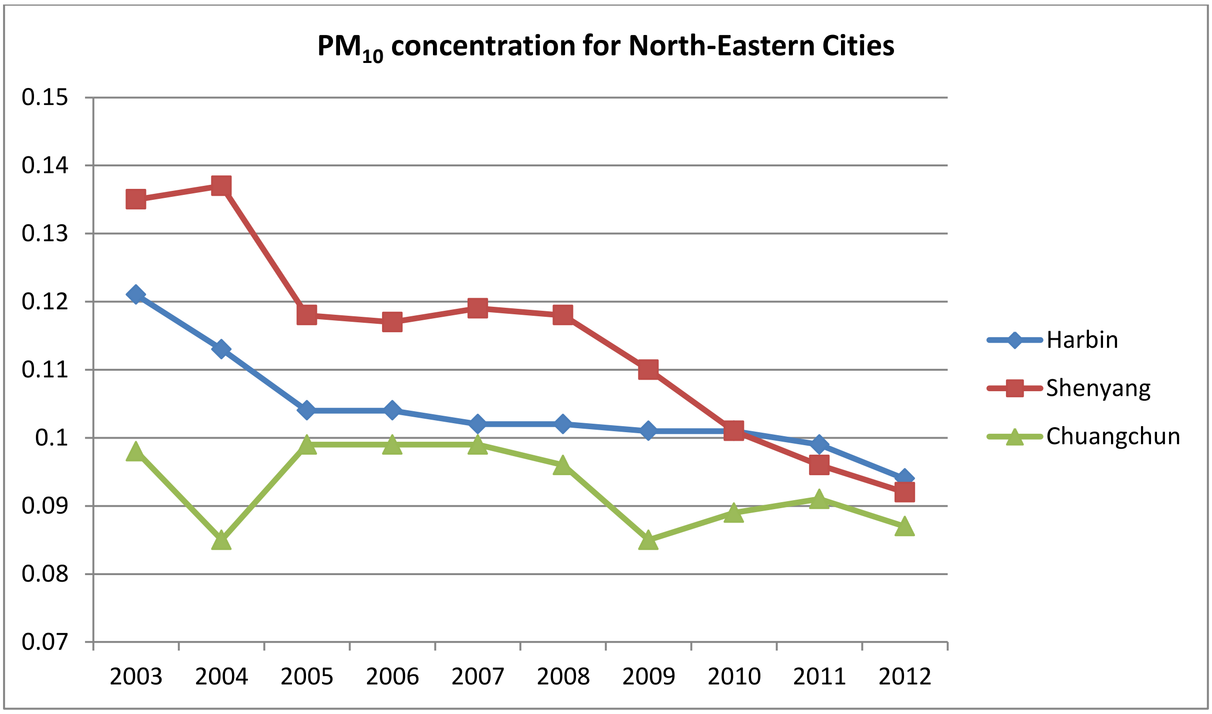


Figure 6.2- D

Sulphur Dioxide (SO_2)

Sulphur dioxide is produced when fuel containing sulphur is burned (mainly coal and oil), and during the smelting of metal and industrial processes. The exposure to high concentration of SO_2 is associated with "breathing, respiratory illness, alterations in pulmonary defenses, and aggravation of existing cardiovascular disease" (Environmental Protection Agent, EPA). The WHO guidelines for SO_2 are 40 μg/m^3 annual mean, 200 μg/m^3 1-hour mean.

Figure 6.3 A-D shows the evolution of annual average SO_2 concentration level for each city. Most of the cities' SO_2 concentrations are between 40 to 80 μg/m^3, and the total trend is decreasing. Specifically, within Eastern cities, two coastal cities (Fuzhou and Haikou) have relatively low levels of SO_2 concentration; Shijiazhuang (Hebei Province) has quite a high level of SO_2 concentration in the beginning (around 150 μg/m^3, 2003) and then the value sharply decreased between 2003 and 2006 (down to 40 μg/m^3). In recent years (2006-2012) SO_2 concentration for Shijiazhuang fluctuates between 40 and 60 μg/m^3, and has been more than halved within this decade (from 150 to 60 μg/m^3). Urumqi in the West had a relatively high SO_2 concentration level before 2011 (consistently over 80 μg/m^3) but this has been reduced to 60 μg/m^3 in 2012. This might be because one of the biggest oil fields in China located in Karamay around Urumqi. Cities in North-Eastern areas have relatively low SO_2 concentration levels. Shenyang has the highest SO_2 concentration level among the three cities in the North-Eastern area and the value is consistently above 60 μg/m^3.

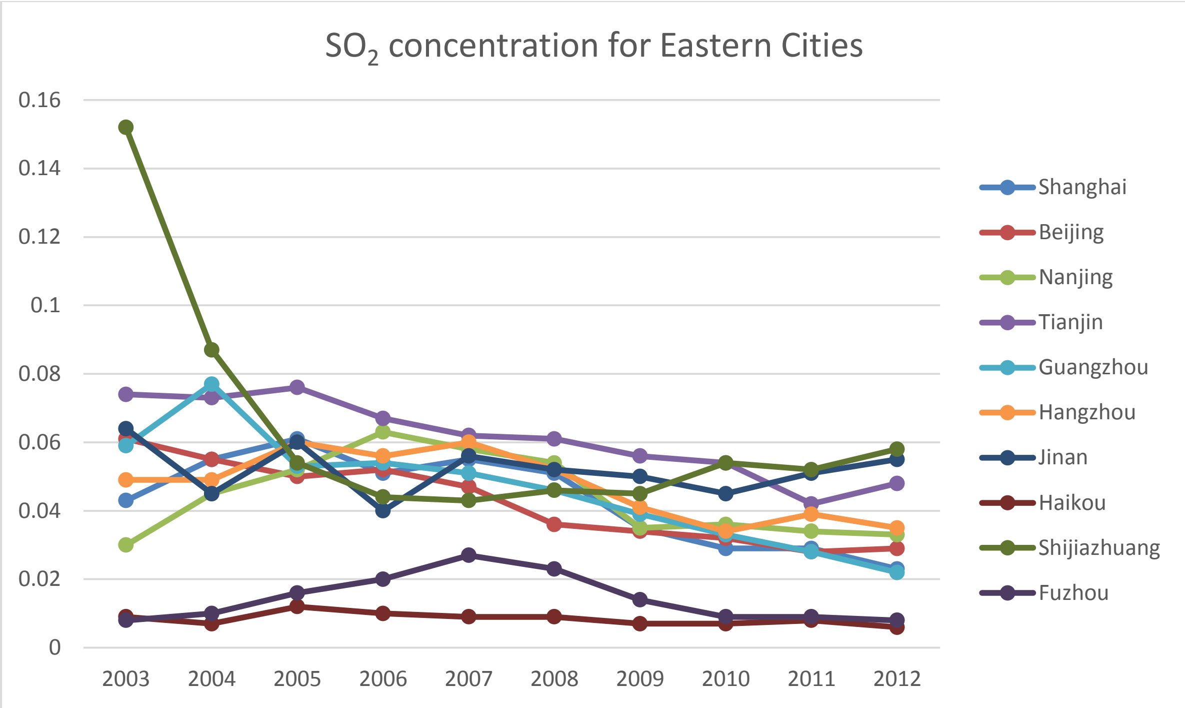


Figure 6.3- A

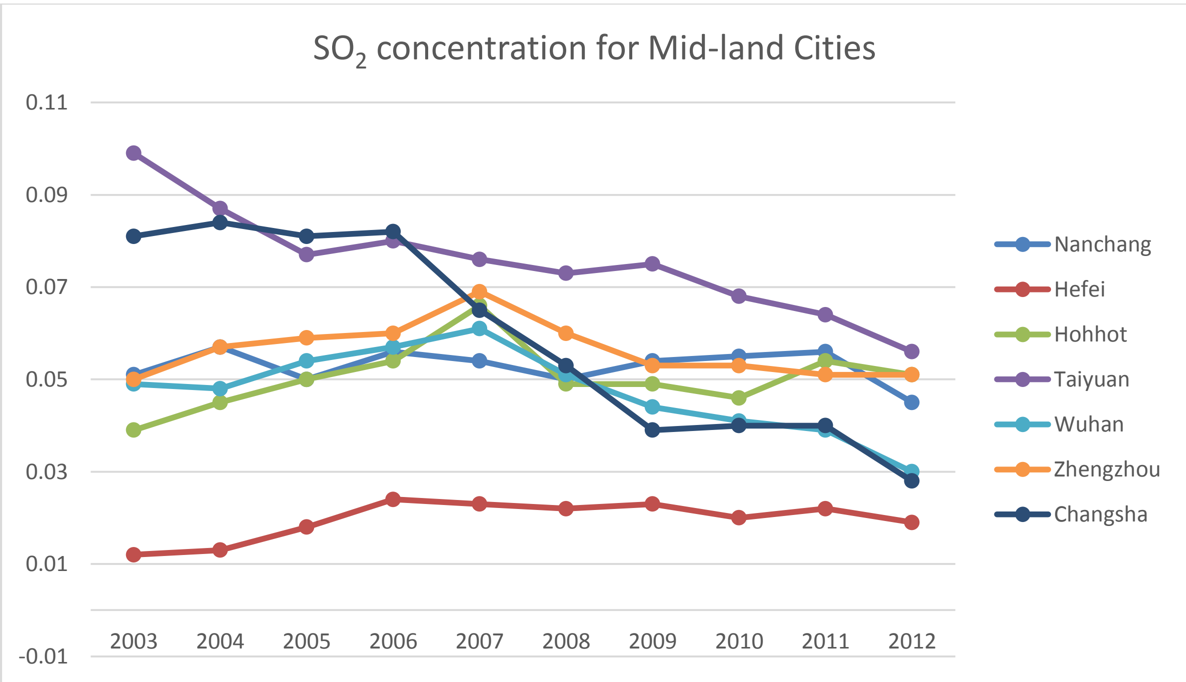


Figure 6.3- B

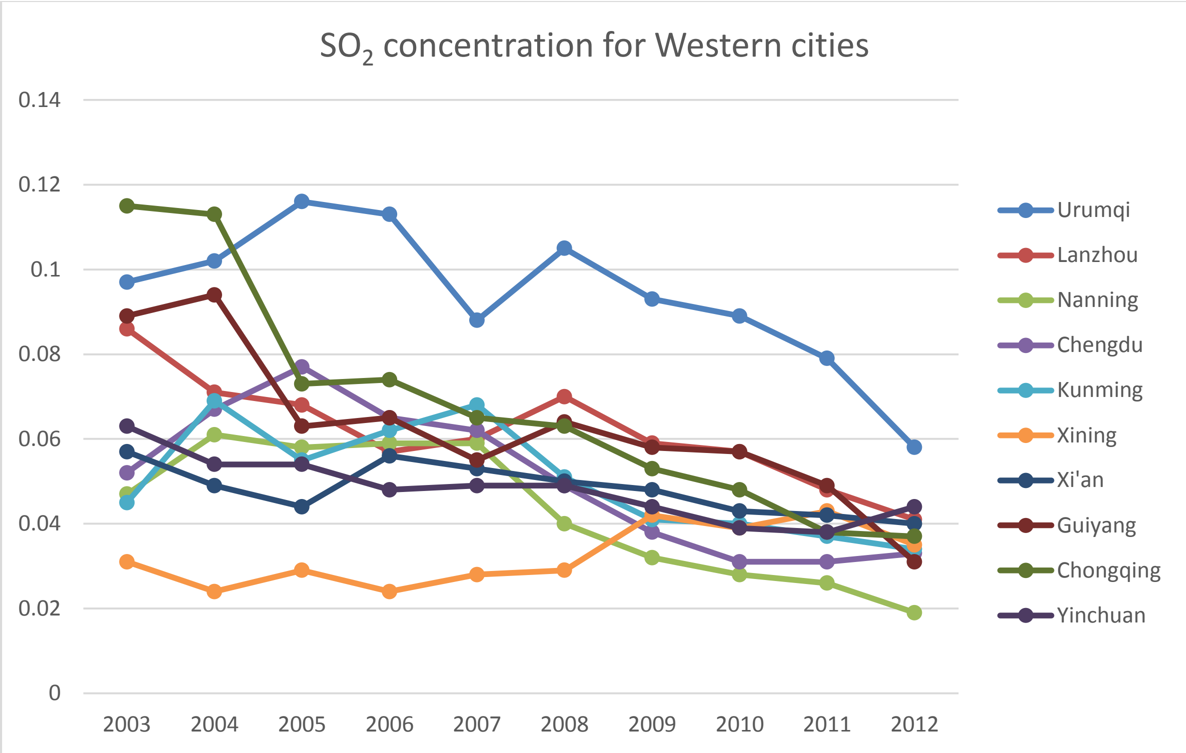

SO_2 concentration for Western cities
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
Urumqi
Lanzhou
Nanning
Chengdu
Kunming
Xining
Xi'an
Guiyang
Chongqing
Yinchuan

Figure 6.3- C

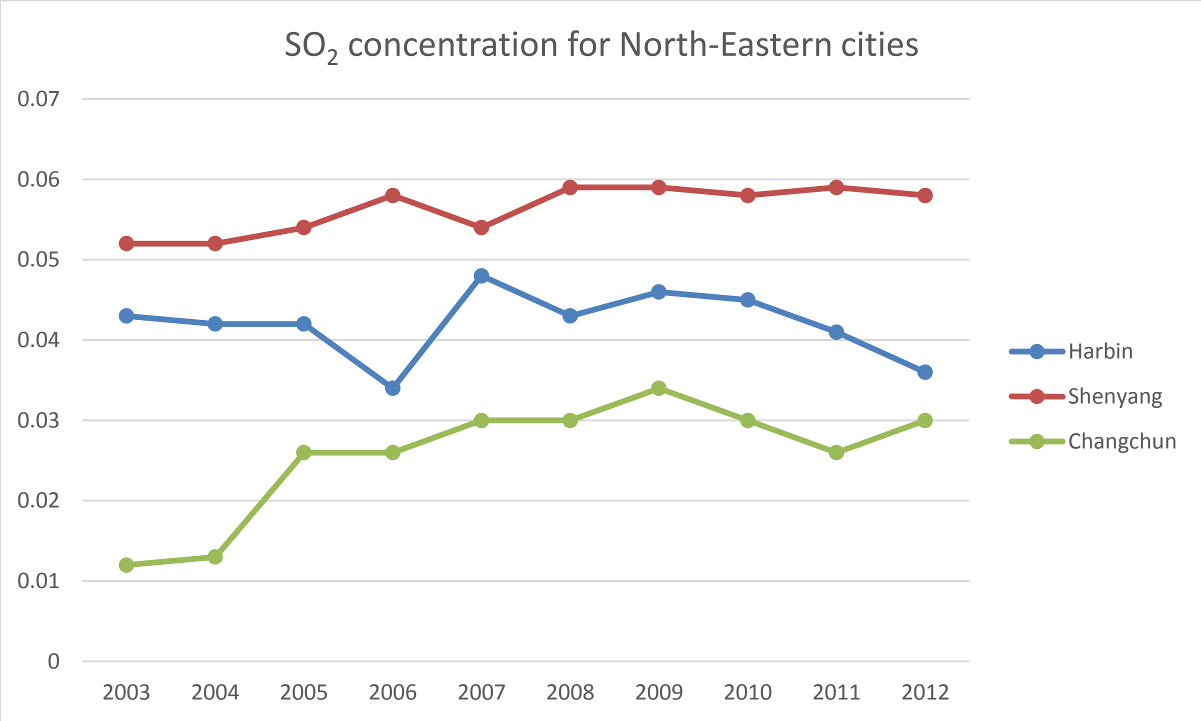


Figure 6.3- D

Nitrogen Dioxide (NO_2)

According to the EPA (Environmental Protection Agency), nitrogen dioxide comes mainly from "motor vehicle exhaust and stationary sources such as electric utilities and industrial boilers", and "it also plays a major role in the atmospheric reactions that produce ground-level ozone (or smog)". In terms of health risk, nitrogen dioxide can irritate the lungs and lower resistance to respiratory infections such as influenza. In addition, nitrogen oxides in the air can significantly contribute to a number of environmental effects such as acid rain and eutrophication in coastal waters. The WHO guidelines for NO2 are 40 μg/m^3 annual mean or 200 μg/m^3 1-hour mean.

Within Eastern cities, Guangzhou is notorious for the acid rain in China as shown in Figure 6.4-A. From 2003 to 2007, the NO_2 concentration level was around 70 μg/m^3. This might reflect the rapid economic development in the Guangzhou region during this time period. However, from 2007 to 2012 the NO_2 concentration level decreased (from around 70 to 50 μg/m^3) which might reflect the effects of environmental regulations.

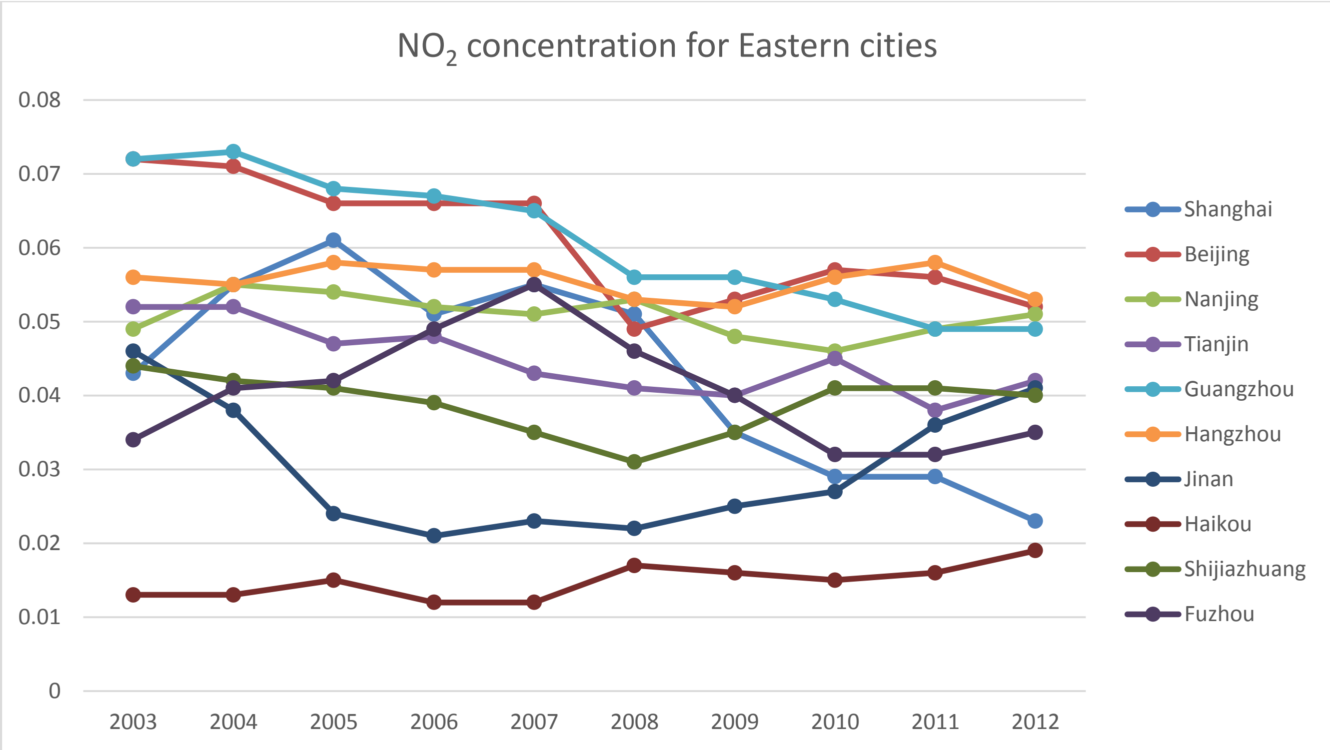


Figure 6.4- A

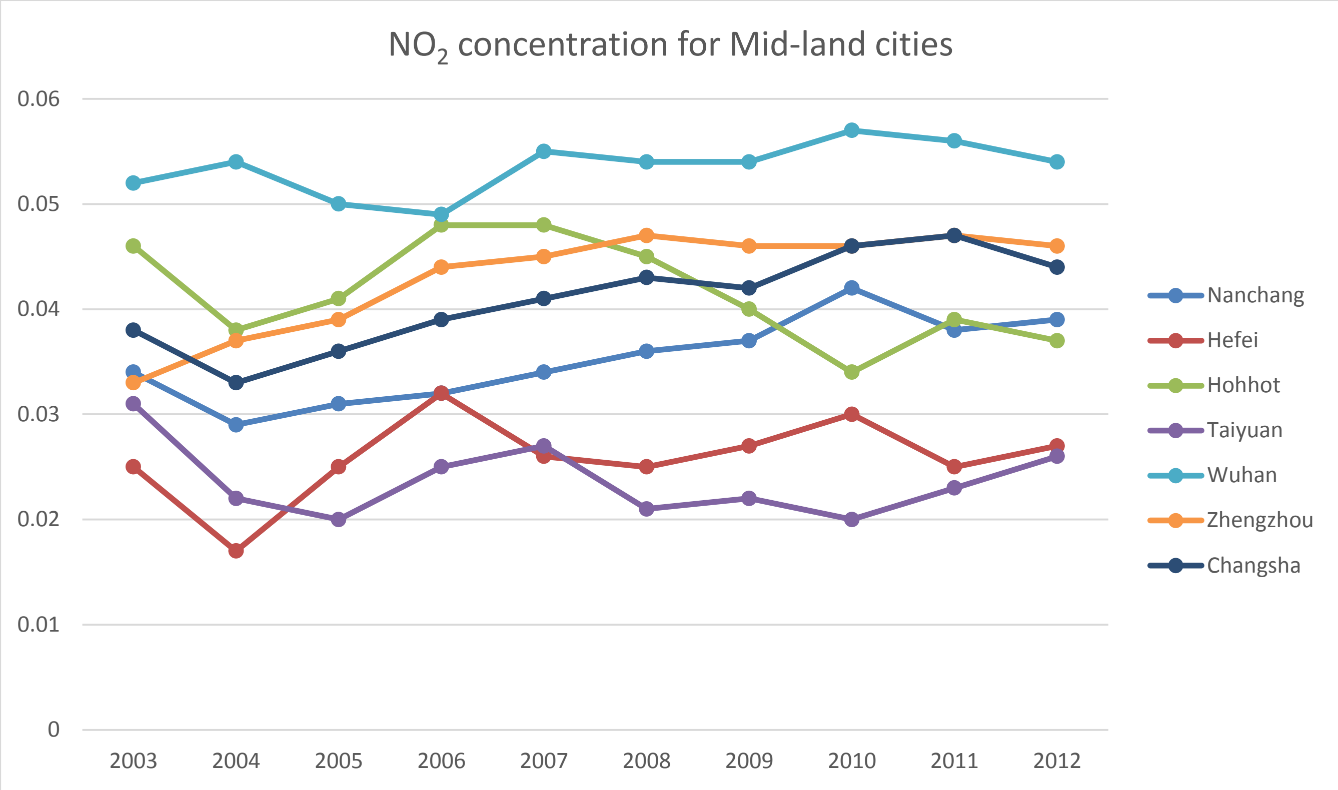


Figure 6.4- B

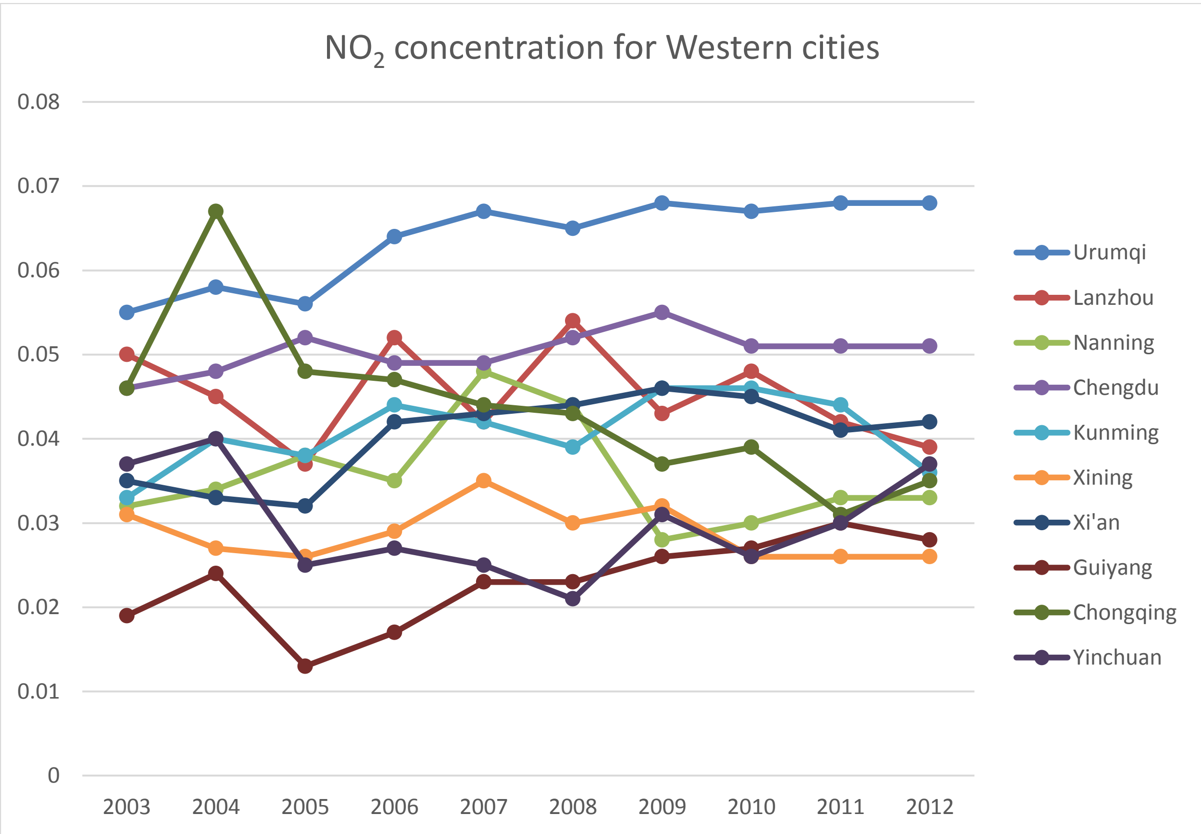


Figure 6.4- C

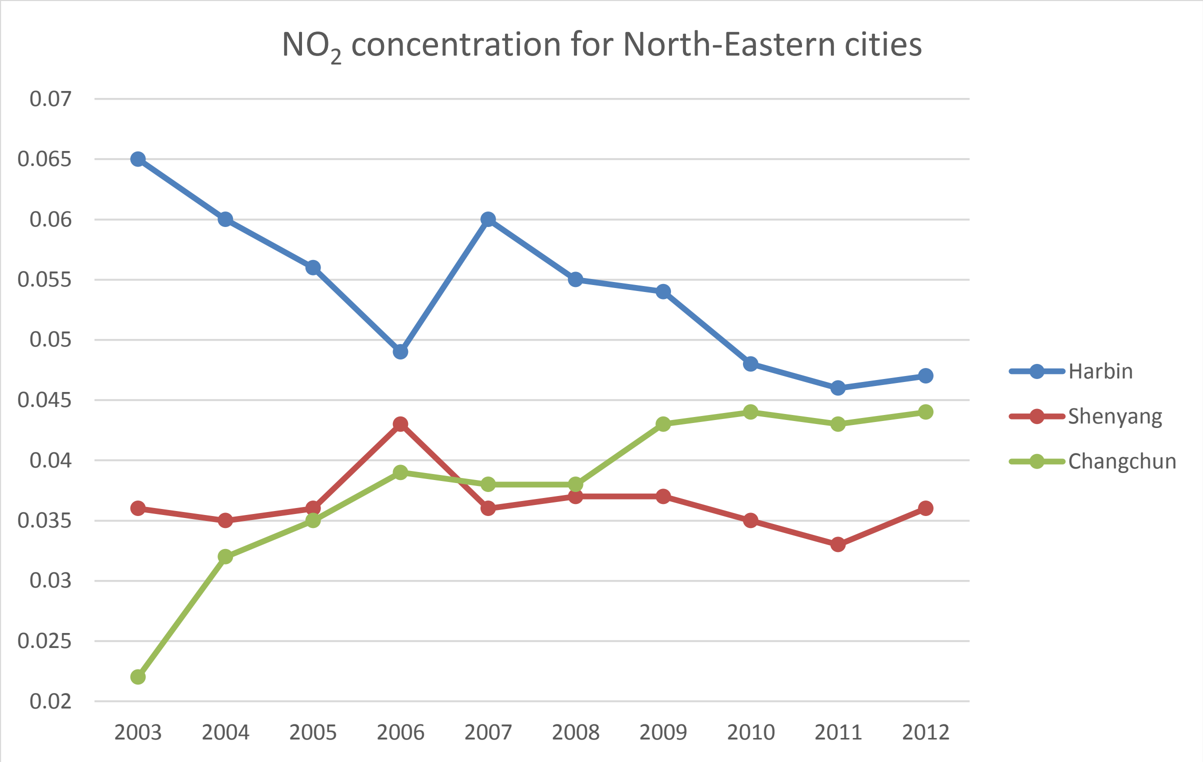


Figure 6.4- D

6.3.2 Empirical Framework

The well-known Environmental Kuznets Curve explores the relationship between environment degradation and per capita income, however, other factors of urban growth may also affect the environment significantly, for example city size. This chapter therefore provides the first comprehensive analysis of the impact of city size on the environment in China. We also use different indicators to represent the size of a city; population size, urban area size, economic size, and the energy consumption size.

Kahn (2006) argues that from an environmental perspective, megacities (with a population over 10 million) are able to provide sustainable advantages. Firstly, megacities can take the advantage of scale of economy and capitalizing in investing on "green" projects, such as "public transit, sewers, and water systems". Secondly, megacities can offer "diverse local labour markets". Small cities tend to specialize in one industry, while large cities are able to construct diversified local industries. Kahn argues that this raises the participation of women in employment market and results in smaller families (Ofek and Merrill 1997; Costa and Kahn 2000). Studies show a significant relationship between urbanisation and population growth (Khan 2006, Arzaghi and Henderson, 2005). And environmentalists encourage of controlling the population growth in order to obtain a long-run sustainability.

Therefore, in line with the empirical framework of Grossman and Krueger (1995), we study the relationship between city size and air pollution by estimating several reduced-form equations. Dependent variable is the three air pollutant indicators respectively, PM_{10}, SO_2, and NO_2; independent variables are city size and other covariates within a flexible function.

Recall that Grossman and Krueger (1995) estimate the pollution on income per capita using the following equation:

$$Y_{it} = \beta_1 G_{it} + \beta_2 {G_{it}}^2 + \beta_3 {G_{it}}^3 + \beta_4 \bar{G}_{it-} + \beta_5 {\bar{G}_{it-}}^2 + \beta_6 {\bar{G}_{it-}}^3 + \beta_7 X_{it} + \epsilon_{it}$$

Y_{it} in their estimation refers to the water or air pollution in station i at year t, G_{it} is GDP per capita in year t in the location where station i is located, $\bar{G}_{it-}$ is the average GDP per capita over previous three years. The dependent variable (air pollution) refers to the annual median of daily concentration at each monitoring site. ${X_{it}}'$ is a vector of other covariates capturing the city characteristics.

We modify Grossman and Krueger (1995)'s model by estimating the following equation (1):

$$Air\ Pollution_{it} = \beta_1 City\ Size_{it} + \beta_2 City\ Size_{it}{}^2 + \beta_3 X_{it}{}' + \epsilon_{it} \qquad (1)$$

where $Air\ Pollution_{it}$ is the air pollution concentration in city i at year t, specifically we estimate three air pollution indicators separately, the PM_{10}, SO_2, and NO_2 annual concentration. The term $City\ Size_{it}$ is the size of a city i at year t, representing by city population size, urban area size, total GDP size (economy size), and energy consumption size respectively to capture the scale effect. $X_{it}{}'$ is a vector of other covariates capturing the city characteristics, including city income level ($GDPPC_{it}$) to capture the wealth level of a city; the industrial composition ($Composition_{it}$) which we use the percentage of primary industry, secondary industry and tertiary industry on total GDP (primary industry has been omitted in the regression result) to capture the composition effect; and the city climate condition ($Climate_{it}$) to control for the effect of climate on cities' air quality.

In addition we also add the output composition of different sources of firms ($domestic$, HMT , $Foreign$). Specifically, the data enables us to split the output from mainland domestic firms, output of Hong Kong, Macau, and Taiwan (HMT) firms and the output produced by foreign firms. We use the percentage shares of these outputs in total GDP in the regression in order to attempt to capture the technique effect. Some studies claim that "foreign multinational firms are less pollution intensive than their domestic counterparts in developing countries" (Eskeland and Harrison, 2003; Cole, Elliott, and Strobl, 2008)[59]. Thus, the estimation equation is:

$$Air\ Pollution_{it} = \beta_1 City\ Size_{it} + \beta_2 City\ Size_{it}{}^2 + \beta_3 GDPPC_{it} + \beta_4 GDPPC_{it}{}^2 + \beta_5 Composition_{it} + \beta_6 Climate_{it} + \beta_7 domestic_{it} + \beta_8 HMT_{it} + \beta_9 Foreign_{it} + \epsilon_{it} \quad (2)$$

The novel feature of our estimation is the attempt of including the scale effect ($City\ Size$), the composition effect ($Composition$) and the technique effect ($domestic$, HMT, $Foreign$). As mentioned in the literature review firstly, we attempt to capture the scale effect by city size, in terms of population, urban area, total GDP, and the energy consumption. Therefore, we estimate each air pollutant by four estimation equations as following:

$$Air\ Pollution_{it} = \beta_1 Population\ Size_{it} + \beta_2 Population\ Size_{it}{}^2 + \beta_3 GDPPC_{it} + \beta_4 GDPPC_{it}{}^2 + \beta_5 Composition_{it} + \beta_6 Climate_{it} + \beta_7 domestic_{it} + \beta_8 HMT_{it} + \beta_9 Foreign_{it} + \epsilon_{it} \quad (3)$$

$$Air\ Pollution_{it} = \beta_1 Urban\ area_{it} + \beta_2 Urban\ area_{it}{}^2 + \beta_3 GDPPC_{it} + \beta_4 GDPPC_{it}{}^2 + \beta_5 Composition_{it} + \beta_6 Climate_{it} + \beta_7 domestic_{it} + \beta_8 HMT_{it} + \beta_9 Foreign_{it} + \epsilon_{it} \quad (4)$$

[59] Although there is no such evidence found in Indonesia by Pargal and Wheeler (1996).

$$Air\ Pollution_{it} = \beta_1 Total\ GDP\ Size_{it} + \beta_2 Total\ GDP\ Size_{it}{}^2 + \beta_3 GDPPC_{it} + \beta_4 GDPPC_{it}{}^2 + \beta_5 Composition_{it} + \beta_6 Climate_{it} + \beta_7 domestic_{it} + \beta_8 HMT_{it} + \beta_9 Foreign_{it} + \epsilon_{it} \quad (5)$$

$$Air\ Pollution_{it} = \beta_1 Energy\ consumption_{it} + \beta_2 Energy\ consumption_{it}{}^2 + \beta_3 GDPPC_{it} + \beta_4 GDPPC_{it}{}^2 + \beta_5 Composition_{it} + \beta_6 Climate_{it} + \beta_7 domestic_{it} + \beta_8 HMT_{it} + \beta_9 Foreign_{it} + \epsilon_{it} \quad (6)$$

Equations (3) to (6) are estimated using both fixed and random effects specifications for three dependent variables as air pollution, PM_{10}, SO_2, NO_2, separately. We rely on city fixed effects (α_i) to capture effects which are specific to each city but have not changed over time, and year specific effects (δ_t) to capture effects which are common to all cities but have changed over time, as the estimation equation (7):

$$Air\ Pollution_{it} = \alpha_i + \delta_t + \beta_1 City\ Size_{it} + \beta_2 City\ Size_{it}{}^2 + \beta_3 GDPPC_{it} + \beta_4 GDPPC_{it}{}^2 + \beta_5 Composition_{it} + \beta_6 Climate_{it} + \beta_7 domestic_{it} + \beta_8 HMT_{it} + \beta_9 Foreign_{it} + \epsilon_{it} \quad (7)$$

We also perform the Hausman test and the Breusch-Pagan Lagrange Multiplier test (LM test). The Hausman test is useful to decide whether fixed or random effect is suitable for the estimation. Fixed effect assumes that some differences within each city may influence the dependent variable (air pollution), and fixed effects is able to control for this by absorbing the effect of those time-invariant characteristics into the intercepts. While random effects capture the influence of differences across cities on the dependent variable, it assumes the variation across cities is random and uncorrelated with the independent variables in the model.

Hausman test is able to test for whether the unobserved omitted time-invariant variables are correlated with the independent variables in the regression, in order to test for whether we can use the random effects or fixed effects. The null hypothesis is that the unique errors (U_i) are not correlated with the explanatory variables and we can perform random effects estimation to capture the influence of differences across cities on dependent variable (Green, 2008, chapter 9).

The Breusch-Pagan Lagrange Multiplier test (LM test) tests between random effects and OLS estimation, the null hypothesis is that variance across individuals is zero, i.e. there are no significant differences across cities in our data. If the null hypothesis cannot be rejected, then random effects is not appropriate. The results shown in Table 6.4-6.6 all reject the null hypothesis, which indicates that there is significant difference across cities in our data.

It should be noted that endogeneity is a potential problem within the estimation equation. The dependent variable - air pollution - potentially affects the city size (population size, urban area, total GDP size and energy consumption). Kahn (2006) argues that within a country with many cities, if the environment deteriorates due to the population growth, this will lead to the moving out of its own residents and impede the potential entrants.

For China, although the migration regulation (*"hukou"* system [60]) has been relaxed in recent years, because of the rapid economic growth and China's Economic Reforms. At this stage (period as our sampled 2003 to 2012), no evidence shows that residences tend to move out of a city because of pollution. Currently the primary motivation for migration in China is to seek better job opportunities and better living amenities in cities, they have not consider the natural environment or pollution level as one of the factors to move out of city. There is no empirical evidence of migration because of lower quality of environment in China. Therefore, we can consider our results are not unduly influenced by endogeneity. Otherwise one can also try system GMM method to control the endogenity and compare the results with this chapter.

6.4 ESTIMATION RESULTS

6.4.1 Results for PM_{10}

We present our estimation results for PM_{10} in Table 6.4 below for both fixed and random effects estimates. [61] The dependent variable is defined as annual average concentrations for each city for PM_{10}, SO_2, NO_2, separately, denoted by PM10, SO2, NO2. All estimates use logarithmic form and we add in the quadratic term sequentially. In each table, estimates (1) and (2) shows the estimates using only population size and its square term respectively; estimates (3) to (6) use the percentage of primary, secondary and tertiary industry of total GDP (percentage of primary industry has been omitted) to control for the impact of industrial

[60] An and Henderson (2006) studies the migration within China and explains that the "hukou" system is an instrument to control the migration across cities and especially between the rural and urban area. A resident's "hukou" is basically registered according to the birth place or mother's legal residence place, which is an "internal passport" in China. The benefits of a legal "hukou" in a city ensure one to access to most of the jobs, housing service and public educations and health care etc.

[61] Econometric software STATA 13.O is used to run the estimations and the major syntaxes include xtreg with fe and re robust; xi: reg varilist i.city i.year to obtain city and year fixed effects.

composition; estimates (7) to (10) include the percentage of output from different sources of firms (domestic, HMT and foreign firms) on total GDP to control for the technique effect of each city.

In addition, we also add the GDP per capita for each city from estimates (3) to (10) to explore the EKC hypothesis as well.

City Size by Population

As shown in Table 6.4_A below, for fixed effects, only when we include both the population size and its square term, the estimated coefficients on population size are significant, which shows the non-linear relationship between population size and the city PM_{10} concentration level. The coefficients for population size are consistently positive, and the coefficients for its square term are consistently negative. These results indicate an inverted U-shape between city population and city PM_{10} concentration levels. This implies that when city population size increases, the PM_{10} concentration will initially increase, however, when the population size passes a certain critical value the PM_{10} concentration will decrease with further population growth.

This might be explained as when initially population size of a city increases, the consumption within a city might increase, e.g. in terms of energy consumption. When population size grows to a certain level, as the urban area tends to be more stable than population, thus higher density of population tends to reduce energy consumption (as people may live in smaller house resulting in less heating, or less commute time resulting in fewer emissions etc.) (Zheng et al., 2011, and Glaeser and Kahn, 2010).

The Hausman test rejects the null hypothesis for most of the estimates which indicates the preferred estimation specification is fixed effect. For estimates (5), (7) and (8) the Hausman test would not run.[62] For model (10), the Hausman test indicates that the preferred estimation uses random effects. We find similar results for estimates of (10) using random effects.

In addition, in the random effect results for PM_{10}, we can find that firstly, the log per capita GDP is consistently negatively significant (when the square term is not included). This suggests that in these 30 major cities in China, when people are getting rich the PM_{10} concentration will decrease. Secondly, results show that the output of domestic firms have a significant positive effect on the concentration of PM_{10}.

[62] Estimates (5), (7), (8) shows that the model fitted on these data fails to meet the asymptotic assumption of the Hausman test.

Table 6.4_A City Size by Population

	Fixed Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10
logpop	0.0330	1.557***	0.0137	1.363**	0.0131	1.398**	0.0457	1.174*	0.0462	1.183*
	(0.0585)	(0.557)	(0.0610)	(0.559)	(0.0612)	(0.564)	(0.0590)	(0.597)	(0.0592)	(0.607)
logpop2		-0.0941***		-0.0836**		-0.0858**		-0.0696*		-0.0702*
		(0.0342)		(0.0344)		(0.0348)		(0.0367)		(0.0373)
loggdppercap			0.00186	0.00140	0.00943	0.0226	0.00182	0.00201	-0.00890	0.00604
			(0.00855)	(0.00847)	(0.0432)	(0.0431)	(0.00880)	(0.00875)	(0.0452)	(0.0457)
loggdppercap2					-0.00109	-0.00305			0.00154	-0.000578
					(0.00608)	(0.00607)			(0.00637)	(0.00643)
secondarygdp			0.000458	0.000464	0.000461	0.000472				
			(0.000342)	(0.000339)	(0.000343)	(0.000340)				
servicegdp			-9.42e-05	-0.000114	-9.49e-05	-0.000116				
			(0.000331)	(0.000328)	(0.000332)	(0.000329)				
logmaxtemp			-0.0179	-0.0204	-0.0179	-0.0205	-0.0223	-0.0233	-0.0224	-0.0233
			(0.0163)	(0.0162)	(0.0163)	(0.0162)	(0.0167)	(0.0166)	(0.0167)	(0.0166)
logmintemp			0.00657	0.00571	0.00652	0.00554	0.00637	0.00581	0.00649	0.00575
			(0.00842)	(0.00835)	(0.00844)	(0.00837)	(0.00847)	(0.00843)	(0.00850)	(0.00847)
logavgpre			-0.0125	-0.00982	-0.0126	-0.0100	-0.0139*	-0.0116	-0.0137*	-0.0117
			(0.00761)	(0.00762)	(0.00764)	(0.00764)	(0.00771)	(0.00776)	(0.00776)	(0.00779)
logavghumi			0.0676***	0.0624***	0.0676***	0.0622***	0.0697***	0.0636***	0.0699***	0.0635***
			(0.0218)	(0.0217)	(0.0218)	(0.0217)	(0.0223)	(0.0224)	(0.0224)	(0.0225)
logavgsun			-0.00495	-0.00440	-0.00478	-0.00391	-0.00383	-0.00279	-0.00413	-0.00267
			(0.00903)	(0.00895)	(0.00910)	(0.00901)	(0.00916)	(0.00913)	(0.00927)	(0.00925)
domestic							0.000142	0.000182	0.000122	0.000189
							(0.000356)	(0.000355)	(0.000366)	(0.000365)
hmt							-0.000674	-0.000478	-0.000679	-0.000475
							(0.000422)	(0.000432)	(0.000423)	(0.000435)
foreign							0.000233	0.000250	0.000252	0.000243
							(0.000562)	(0.000559)	(0.000569)	(0.000566)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-2.184***	-8.196***	-2.030***	-7.320***	-2.035***	-7.479***	-2.304***	-6.763***	-2.293***	-6.804***
	(0.542)	(2.248)	(0.592)	(2.257)	(0.594)	(2.282)	(0.548)	(2.409)	(0.551)	(2.458)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.895	0.898	0.903	0.905	0.903	0.905	0.900	0.902	0.900	0.902
Hausman test	0.0000	0.0012	0.0000	0.0000	-	0.0000	-	-	0.0465	0.9989

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

	Random Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10
logpop	-0.151***	0.319	-0.0890**	0.557	-0.0420	1.035*	0.0214	1.137*	0.0468	1.490***
	(0.0460)	(0.623)	(0.0417)	(0.560)	(0.0425)	(0.548)	(0.0408)	(0.585)	(0.0406)	(0.568)
logpop2		-0.0292		-0.0394		-0.0658**		-0.0664*		-0.0867**
		(0.0383)		(0.0344)		(0.0336)		(0.0356)		(0.0345)
loggdppercap			-0.0283***	-0.0294***	0.160***	0.171***	-0.0329***	-0.0336***	0.201***	0.217***
			(0.00890)	(0.00897)	(0.0402)	(0.0406)	(0.0103)	(0.0106)	(0.0440)	(0.0456)
loggdppercap2					-0.0255***	-0.0272***			-0.0318***	-0.0340***
					(0.00533)	(0.00540)			(0.00584)	(0.00605)
secondarygdp			0.000363	0.000364	0.000524	0.000536				
			(0.000381)	(0.000383)	(0.000366)	(0.000367)				
servicegdp			-0.000796**	-0.000806**	-0.000484	-0.000481				
			(0.000356)	(0.000358)	(0.000347)	(0.000347)				
logmaxtemp			-0.00834	-0.00861	-0.0106	-0.0116	-0.00907	-0.00424	-0.0107	-0.00578
			(0.0176)	(0.0177)	(0.0169)	(0.0169)	(0.0191)	(0.0193)	(0.0181)	(0.0183)
logmintemp			-0.00380	-0.00474	-0.00202	-0.00319	-0.00619	-0.00946	-0.00574	-0.00888
			(0.00942)	(0.00949)	(0.00901)	(0.00904)	(0.0105)	(0.0108)	(0.00984)	(0.0101)
logavgpre			-0.00820	-0.00672	-0.0116	-0.00945	-0.00853	-0.00541	-0.0129	-0.00951
			(0.00858)	(0.00869)	(0.00823)	(0.00829)	(0.00952)	(0.00986)	(0.00898)	(0.00926)
logavghumi			0.0540**	0.0492**	0.0522**	0.0455**	0.0305	0.00748	0.0292	0.00472
			(0.0225)	(0.0227)	(0.0216)	(0.0217)	(0.0244)	(0.0248)	(0.0232)	(0.0236)
logavgsun			-0.00532	-0.00434	-0.000803	0.000716	0.00825	0.0140	0.0120	0.0179*
			(0.00997)	(0.0100)	(0.00961)	(0.00962)	(0.0109)	(0.0112)	(0.0104)	(0.0106)
domestic							0.00108**	0.00120***	0.00118***	0.00132***
							(0.000446)	(0.000464)	(0.000419)	(0.000434)
hmt							-6.95e-05	0.000250	-0.000162	0.000204
							(0.000533)	(0.000564)	(0.000500)	(0.000527)
foreign							0.00108	0.00108	0.000257	0.000223
							(0.000700)	(0.000727)	(0.000676)	(0.000697)
Constant	-1.088***	-2.967	-1.444***	-4.056*	-2.186***	-6.558***	-2.389***	-7.011***	-2.939***	-8.903***
	(0.376)	(2.523)	(0.359)	(2.267)	(0.390)	(2.241)	(0.344)	(2.398)	(0.356)	(2.340)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

The graph below also suggests an inverted-U relationship between population size and PM_{10} concentrations.

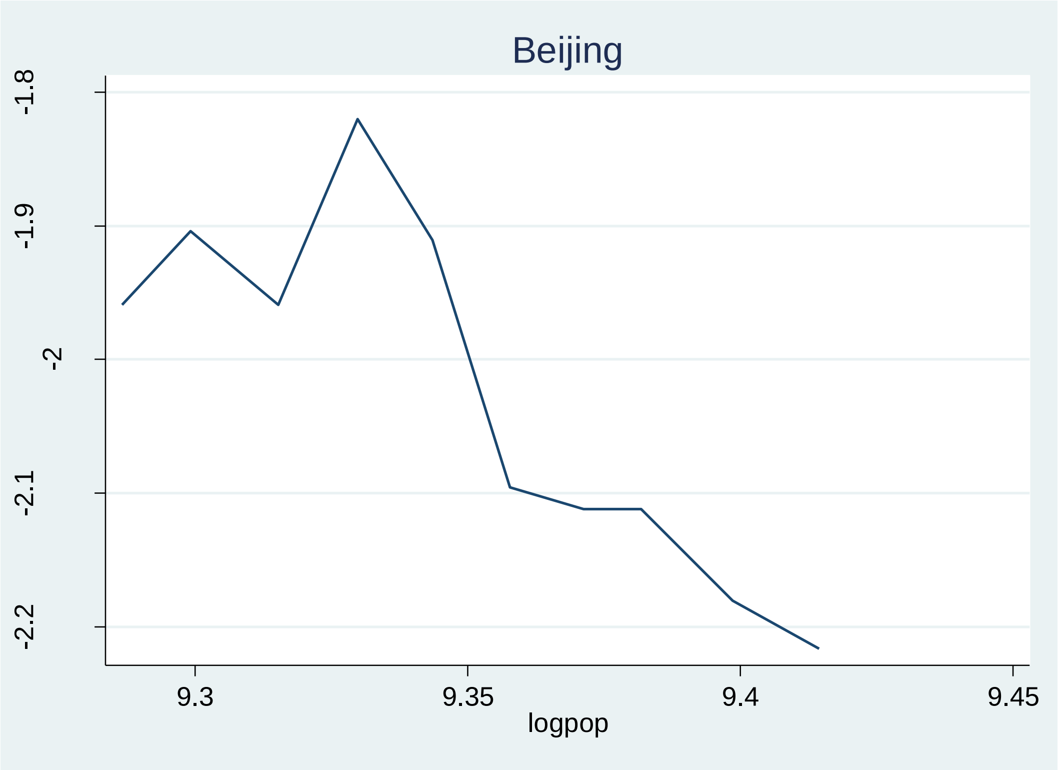


Figure 6.5: plots PM_{10} concentration level against population for Beijing.

City Size by Urban Area

Table 6.4_B shows the results for urban area as the city size in estimation equation (4). For estimates (6) and (10), including all of the independent variables (both urban area, GDP per capita and their square terms) for composition and technique effect respectively, the Hausman test suggests that the random effect specification is preferred. In the random effect table for estimates (6) and (10), the urban area tends to have an inverted-U relationship with

PM_{10} concentration, similar to the population size. The secondary industry (mainly manufacturing industry) shows a significant positive impact on PM_{10} concentration levels, but the magnitude of the estimated coefficient is quite small (0.000625). And the output of domestic firms shows a significant positive effect on PM_{10} pollution, although again the estimated coefficient on this term is small (0.00105). Since we are estimating in natural logarithms, so the marginal effects can be explained as elasticity, our results show that a 1% increase in a city's composition of domestic firm's output in total GDP leads to a 0.00105% increase in PM_{10} concentration in a city.

The urban area in this table refers to the built-up area. The built-up area captures the city development both in terms of population and economic scale, it includes approved land use by government and the real developed non-agriculture industrial land use in the downtown and the land highly connected with city development scattered in the surrounding suburban areas which have the core facilities (e.g. airport, sewage disposal plant, communication station, etc.).

Therefore, when urban area increases, it represents the development of a city, the PM_{10} concentration will increase accordingly in most models (fixed effects results for estimates (1), (3), (5), (7), (9)). But when urban area reaches a certain size, the PM_{10} concentration might decrease (random effects for estimates (6) and (8)).

Fixed effect results also show that precipitation plays a negative role in PM_{10} concentration level, i.e. the more precipitation in a city, the lower its PM_{10} concentrations (annual average).

Table 6.4_ B City Size by Urban Area (Built-Up Area)

	Fixed Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10
VARIABLES	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10
logbuiltup	0.0170*	-0.00121	0.0155*	0.0203	0.0161*	0.0196	0.0187**	-0.0213	0.0188**	-0.0216
	(0.00905)	(0.0539)	(0.00893)	(0.0539)	(0.00903)	(0.0540)	(0.00908)	(0.0555)	(0.00919)	(0.0557)
logbuiltup2		0.00275		-0.000716		-0.000534		0.00607		0.00613
		(0.00803)		(0.00803)		(0.00805)		(0.00831)		(0.00835)
loggdppercap			0.00174	0.00164	0.0206	0.0204	0.00141	0.00192	0.00434	0.00705
			(0.00850)	(0.00859)	(0.0433)	(0.0435)	(0.00873)	(0.00877)	(0.0453)	(0.0455)
loggdppercap2					-0.00271	-0.00269			-0.000420	-0.000735
					(0.00609)	(0.00611)			(0.00638)	(0.00640)
secondarygdp			0.000485	0.000488	0.000494	0.000496				
			(0.000339)	(0.000341)	(0.000340)	(0.000342)				
servicegdp			-6.20e-05	-6.18e-05	-6.00e-05	-5.98e-05				
			(0.000321)	(0.000322)	(0.000321)	(0.000322)				
logmaxtemp			-0.0174	-0.0174	-0.0176	-0.0176	-0.0214	-0.0213	-0.0214	-0.0213
			(0.0161)	(0.0161)	(0.0161)	(0.0162)	(0.0165)	(0.0165)	(0.0165)	(0.0165)
logmintemp			0.00639	0.00637	0.00624	0.00623	0.00622	0.00623	0.00618	0.00616
			(0.00837)	(0.00838)	(0.00839)	(0.00840)	(0.00839)	(0.00840)	(0.00843)	(0.00843)
logavgpre			-0.0129*	-0.0129*	-0.0131*	-0.0131*	-0.0141*	-0.0139*	-0.0141*	-0.0140*
			(0.00754)	(0.00756)	(0.00758)	(0.00760)	(0.00764)	(0.00765)	(0.00770)	(0.00770)
logavghumi			0.0669***	0.0669***	0.0669***	0.0669***	0.0685***	0.0691***	0.0684***	0.0690***
			(0.0215)	(0.0215)	(0.0215)	(0.0216)	(0.0220)	(0.0220)	(0.0221)	(0.0221)
logavgsun			-0.00554	-0.00551	-0.00511	-0.00509	-0.00508	-0.00549	-0.00499	-0.00535
			(0.00895)	(0.00897)	(0.00902)	(0.00904)	(0.00908)	(0.00911)	(0.00918)	(0.00921)
domestic							0.000139	0.000139	0.000145	0.000148
							(0.000353)	(0.000354)	(0.000363)	(0.000363)
hmt							-0.000793*	-0.000847**	-0.000793*	-0.000846**
							(0.000421)	(0.000428)	(0.000422)	(0.000428)
foreign							0.000226	0.000178	0.000221	0.000168
							(0.000558)	(0.000562)	(0.000565)	(0.000570)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-1.923***	-1.893***	-1.946***	-1.954***	-1.976***	-1.981***	-1.918***	-1.853***	-1.923***	-1.860***
	(0.0433)	(0.0985)	(0.122)	(0.148)	(0.140)	(0.161)	(0.0819)	(0.121)	(0.106)	(0.136)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.896	0.896	0.904	0.904	0.904	0.904	0.902	0.902	0.902	0.902
Hausman test	-	-	0.0000	0.4610	0.0090	0.7295	0.0000	0.0000	0.0160	1.0000

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

	Random Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10
VARIABLES	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10
logbuiltup	-0.00865	0.225***	-0.00129	0.193***	0.0110	0.138**	-0.000170	0.215***	0.0130	0.141**
	(0.0112)	(0.0628)	(0.0102)	(0.0572)	(0.00989)	(0.0566)	(0.0114)	(0.0625)	(0.0109)	(0.0617)
logbuiltup2		-0.0347***		-0.0288***		-0.0190**		-0.0318***		-0.0192**
		(0.00918)		(0.00836)		(0.00836)		(0.00910)		(0.00912)
loggdppercap			-0.0326***	-0.0313***	0.177***	0.154***	-0.0327***	-0.0302***	0.204***	0.177***
			(0.00882)	(0.00871)	(0.0400)	(0.0416)	(0.0105)	(0.0102)	(0.0450)	(0.0465)
loggdppercap2					-0.0283***	-0.0250***			-0.0322***	-0.0284***
					(0.00528)	(0.00550)			(0.00598)	(0.00622)
secondarygdp			0.000395	0.000499	0.000574	0.000625*				
			(0.000385)	(0.000381)	(0.000366)	(0.000369)				
servicegdp			-0.000738**	-0.000622*	-0.000411	-0.000371				
			(0.000359)	(0.000356)	(0.000345)	(0.000348)				
logmaxtemp			-0.0125	-0.0130	-0.0127	-0.0116	-0.00653	-0.00849	-0.00786	-0.00889
			(0.0176)	(0.0173)	(0.0167)	(0.0168)	(0.0192)	(0.0188)	(0.0182)	(0.0181)
logmintemp			-0.00509	-0.00443	-0.00282	-0.00311	-0.00707	-0.00556	-0.00668	-0.00583
			(0.00952)	(0.00939)	(0.00901)	(0.00908)	(0.0107)	(0.0104)	(0.00997)	(0.00992)
logavgpre			-0.00797	-0.00915	-0.0119	-0.0118	-0.00815	-0.0102	-0.0129	-0.0135
			(0.00867)	(0.00856)	(0.00823)	(0.00829)	(0.00967)	(0.00947)	(0.00909)	(0.00904)
logavghumi			0.0564**	0.0541**	0.0532**	0.0483**	0.0247	0.0271	0.0246	0.0260
			(0.0226)	(0.0223)	(0.0216)	(0.0216)	(0.0245)	(0.0240)	(0.0233)	(0.0232)
logavgsun			-0.00291	-0.00227	0.000409	0.00109	0.00891	0.00879	0.0113	0.0110
			(0.0100)	(0.00987)	(0.00953)	(0.00957)	(0.0110)	(0.0108)	(0.0104)	(0.0103)
domestic							0.00106**	0.000914**	0.00115***	0.00105**
							(0.000452)	(0.000444)	(0.000423)	(0.000423)
hmt							-6.75e-05	9.78e-05	-0.000242	-0.000120
							(0.000541)	(0.000531)	(0.000506)	(0.000507)
foreign							0.000933	0.00100	9.94e-05	0.000241
							(0.000700)	(0.000685)	(0.000679)	(0.000677)
Constant	-2.27***	-2.606***	-2.148***	-2.452***	-2.605***	-2.751***	-2.207***	-2.522***	-2.606***	-2.748***
	(0.0698)	(0.110)	(0.142)	(0.165)	(0.159)	(0.173)	(0.105)	(0.136)	(0.123)	(0.139)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

City Size by Total GDP

For the economic size of a city, we estimate the impact of total GDP on the PM_{10} concentration level. The square term of total GDP is consistently significantly negative either for fixed or random effects, which indicates the non-linear relationship between total GDP and PM_{10} concentration.

When we exclude the square term of total GDP, GDP size shows a negative effect on PM_{10} concentrations in the random effects models, which might indicate that the larger the economic scale of a city the lower are its PM_{10} concentrations. These 30 major cities are in the downward sloping part of the inverted-U shape relationship.

Output of secondary industry and domestic firms still play a positive role in PM_{10} concentration, but still with small effects (around 0.00058 and 0.000868 respectively).

Table 6.4_ C City Size by Total GDP

Fixed Effects

VARIABLES	(1) logpm_10	(2) logpm_10	(3) logpm_10	(4) logpm_10	(5) logpm_10	(6) logpm_10	(7) logpm_10	(8) logpm_10	(9) logpm_10	(10) logpm_10
loggdp	-0.0203	0.0370	-0.0758	-0.0134	-0.0753	-0.0136	0.00678	0.109	0.00489	0.107
	(0.0573)	(0.0622)	(0.0671)	(0.0717)	(0.0674)	(0.0718)	(0.0619)	(0.0766)	(0.0628)	(0.0773)
loggdp2		-0.0147**		-0.0152**		-0.0153**		-0.0186**		-0.0186**
		(0.00647)		(0.00651)		(0.00656)		(0.00833)		(0.00835)
loggdppercap			0.00305	-0.00156	0.00799	-0.00739	0.00167	-0.00223	-0.00736	-0.0119
			(0.00860)	(0.00875)	(0.0431)	(0.0432)	(0.00884)	(0.00894)	(0.0457)	(0.0453)
loggdppercap2					-0.000710	0.000832			0.00130	0.00139
					(0.00607)	(0.00605)			(0.00644)	(0.00639)
secondarygdp			0.000487	0.000502	0.000489	0.000500				
			(0.000341)	(0.000338)	(0.000342)	(0.000339)				
servicegdp			-0.000187	-0.000174	-0.000186	-0.000174				
			(0.000328)	(0.000325)	(0.000328)	(0.000325)				
logmaxtemp			-0.0159	-0.0194	-0.0159	-0.0193	-0.0210	-0.0215	-0.0209	-0.0215
			(0.0162)	(0.0161)	(0.0162)	(0.0162)	(0.0166)	(0.0165)	(0.0166)	(0.0165)
logmintemp			0.00633	0.00572	0.00629	0.00576	0.00690	0.00649	0.00701	0.00661
			(0.00840)	(0.00833)	(0.00842)	(0.00835)	(0.00845)	(0.00839)	(0.00849)	(0.00842)
logavgpre			-0.0118	-0.0111	-0.0118	-0.0110	-0.0135*	-0.0126	-0.0133*	-0.0125
			(0.00759)	(0.00752)	(0.00762)	(0.00756)	(0.00771)	(0.00766)	(0.00776)	(0.00771)
logavghumi			0.0608***	0.0620***	0.0608***	0.0620***	0.0683***	0.0666***	0.0684***	0.0666***
			(0.0223)	(0.0221)	(0.0223)	(0.0221)	(0.0227)	(0.0225)	(0.0227)	(0.0226)
logavgsun			-0.00408	-0.00438	-0.00397	-0.00451	-0.00443	-0.00413	-0.00466	-0.00438
			(0.00903)	(0.00895)	(0.00909)	(0.00901)	(0.00920)	(0.00913)	(0.00929)	(0.00922)
domestic							0.000149	0.000458	0.000132	0.000440
							(0.000358)	(0.000381)	(0.000369)	(0.000391)
hmt							-0.000685	-0.000350	-0.000691	-0.000357
							(0.000429)	(0.000452)	(0.000431)	(0.000454)
foreign							0.000236	0.000539	0.000252	0.000556
							(0.000563)	(0.000575)	(0.000569)	(0.000581)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-1.804***	-1.772***	-1.606***	-1.556***	-1.615***	-1.545***	-1.910***	-1.977***	-1.889***	-1.955***
	(0.216)	(0.215)	(0.287)	(0.285)	(0.298)	(0.297)	(0.246)	(0.246)	(0.266)	(0.266)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.895	0.897	0.903	0.905	0.903	0.905	0.900	0.902	0.900	0.902
Hausman test	0.0238	-	0.1385	-	0.8767	-	0.0009	-	-	0.0000

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

Random Effects

VARIABLES	(1) logpm_10	(2) logpm_10	(3) logpm_10	(4) logpm_10	(5) logpm_10	(6) logpm_10	(7) logpm_10	(8) logpm_10	(9) logpm_10	(10) logpm_10
loggdp	-0.147***	-0.0779**	-0.118***	-0.0448	-0.110***	-0.0414	-0.127***	0.00387	-0.110***	0.0218
	(0.0121)	(0.0363)	(0.0146)	(0.0395)	(0.0187)	(0.0401)	(0.0176)	(0.0626)	(0.0226)	(0.0635)
loggdp2		-0.0129**		-0.0131**		-0.0127*		-0.0197**		-0.0195**
		(0.00640)		(0.00657)		(0.00662)		(0.00941)		(0.00932)
loggdppercap			-0.00119	-0.00566	0.0298	0.0154	-0.00798	-0.0139	0.0550	0.0591
			(0.00868)	(0.00892)	(0.0436)	(0.0440)	(0.00973)	(0.0102)	(0.0500)	(0.0510)
loggdppercap2					-0.00444	-0.00301			-0.00898	-0.0104
					(0.00615)	(0.00617)			(0.00703)	(0.00715)
secondarygdp			0.000569*	0.000590*	0.000580*	0.000597*				
			(0.000342)	(0.000340)	(0.000342)	(0.000341)				
servicegdp			-0.000290	-0.000246	-0.000271	-0.000235				
			(0.000323)	(0.000321)	(0.000323)	(0.000322)				
logmaxtemp			-0.00699	-0.0111	-0.00770	-0.0114	-0.0104	-0.00877	-0.0109	-0.00947
			(0.0158)	(0.0158)	(0.0158)	(0.0158)	(0.0173)	(0.0175)	(0.0172)	(0.0174)
logmintemp			0.00133	0.00106	0.00141	0.00111	-0.00221	-0.00408	-0.00230	-0.00409
			(0.00845)	(0.00840)	(0.00843)	(0.00839)	(0.00936)	(0.00959)	(0.00930)	(0.00950)
logavgpre			-0.0104	-0.00978	-0.0109	-0.0101	-0.0110	-0.00967	-0.0120	-0.0108
			(0.00769)	(0.00765)	(0.00770)	(0.00767)	(0.00853)	(0.00873)	(0.00850)	(0.00868)
logavghumi			0.0442**	0.0474**	0.0452**	0.0478**	0.0328	0.0254	0.0333	0.0261
			(0.0205)	(0.0204)	(0.0205)	(0.0204)	(0.0223)	(0.0225)	(0.0222)	(0.0224)
logavgsun			-0.00436	-0.00479	-0.00378	-0.00436	0.00213	0.00395	0.00351	0.00551
			(0.00892)	(0.00886)	(0.00895)	(0.00892)	(0.00981)	(0.00999)	(0.00983)	(0.0100)
domestic							0.000351	0.000741	0.000466	0.000868*
							(0.000407)	(0.000451)	(0.000415)	(0.000457)
hmt							-0.000678	-0.000273	-0.000639	-0.000234
							(0.000479)	(0.000519)	(0.000477)	(0.000515)
foreign							-6.18e-05	0.000201	-0.000145	0.000105
							(0.000639)	(0.000674)	(0.000638)	(0.000672)
Constant	-1.909***	-1.987***	-2.064***	-2.132***	-2.134***	-2.178***	-1.925***	-2.095***	-2.063***	-2.252***
	(0.0602)	(0.0719)	(0.123)	(0.127)	(0.157)	(0.158)	(0.0970)	(0.126)	(0.145)	(0.166)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

City Size by Energy Consumption

Coal gas consumption shows a significant inverted-U shape with PM_{10} concentration either in fixed or random effects. For fixed effects which can absorb the effect of those time-invariant characteristics, coal and gas consumption shows a significant inverted-U shape when only energy consumption is included in the regression equation (regression (2) in Table 6.4_D fixed effects). Also for the random effects in regression (6) and (10) in Table 6.4_D random effects, when all independent variables are included for industrial composition effect (regression (6)) and technique effect (regression (10)), we find that coal and gas consumption has a significant impact on the PM_{10} concentrations of a city.

In addition, the EKC hypothesis seems to held in the random effects models in every specification (Table 6.4_D_random effects). These show an inverted-U shape between GDP per capita and PM_{10} concentrations.

In the random effects models, output from domestic firms shows a significant positive effect on PM_{10} concentration. This shows that secondary industry plays a positive role in PM_{10} concentration, in this section, we find that tertiary industry shows a negative impact on PM_{10} concentration, i.e. the more service industry in a city's industry composition, the lower are PM_{10} concentration levels.

Table 6.4_ D City Size by Energy Consumption

	Fixed Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10
logelec	-0.0115*	9.27e-05	-0.0101	0.0154	-0.0103	0.0151	-0.0103	0.00814	-0.0103	0.00807
	(0.00625)	(0.0292)	(0.00638)	(0.0293)	(0.00641)	(0.0294)	(0.00650)	(0.0300)	(0.00653)	(0.0301)
logelec2		-0.00149		-0.00342		-0.00340		-0.00242		-0.00242
		(0.00395)		(0.00398)		(0.00398)		(0.00406)		(0.00407)
logcoalgas	0.000420	0.0365*	0.00138	0.0293	0.00137	0.0316	0.000339	0.0312	0.000350	0.0319
	(0.00548)	(0.0212)	(0.00547)	(0.0214)	(0.00548)	(0.0218)	(0.00555)	(0.0217)	(0.00556)	(0.0221)
logcoalgas2		-0.00569*		-0.00445		-0.00481		-0.00491		-0.00503
		(0.00321)		(0.00324)		(0.00331)		(0.00329)		(0.00335)
loglpg	0.000365	0.000493	-0.00105	-0.00606	-0.00113	-0.00533	-0.00145	0.00186	-0.00144	0.00192
	(0.00678)	(0.0287)	(0.00672)	(0.0284)	(0.00674)	(0.0285)	(0.00692)	(0.0292)	(0.00694)	(0.0293)
loglpg2		-0.000351		0.000436		0.000291		-0.000779		-0.000798
		(0.00398)		(0.00394)		(0.00395)		(0.00406)		(0.00407)
loggdppercap			0.00377	0.00251	0.0160	0.0268	0.00394	0.00222	-0.000552	0.0109
			(0.00867)	(0.00889)	(0.0434)	(0.0441)	(0.00892)	(0.00912)	(0.0455)	(0.0461)
loggdppercap2					-0.00175	-0.00350			0.000643	-0.00125
					(0.00608)	(0.00622)			(0.00638)	(0.00651)
secondarygdp			0.000468	0.000506	0.000473	0.000516				
			(0.000341)	(0.000344)	(0.000342)	(0.000345)				
servicegdp			-8.40e-05	-5.66e-05	-8.38e-05	-5.44e-05				
			(0.000322)	(0.000323)	(0.000323)	(0.000323)				
logmaxtemp			-0.0208	-0.0191	-0.0209	-0.0192	-0.0237	-0.0217	-0.0237	-0.0216
			(0.0164)	(0.0166)	(0.0165)	(0.0166)	(0.0169)	(0.0170)	(0.0169)	(0.0170)
logmintemp			0.00785	0.00710	0.00776	0.00692	0.00811	0.00740	0.00816	0.00730
			(0.00845)	(0.00848)	(0.00847)	(0.00850)	(0.00849)	(0.00854)	(0.00852)	(0.00858)
logavgpre			-0.0101	-0.0105	-0.0103	-0.0108	-0.0112	-0.0116	-0.0111	-0.0117
			(0.00770)	(0.00772)	(0.00773)	(0.00775)	(0.00783)	(0.00784)	(0.00787)	(0.00788)
logavghumi			0.0674***	0.0642***	0.0674***	0.0638***	0.0679***	0.0643***	0.0680***	0.0640***
			(0.0216)	(0.0219)	(0.0217)	(0.0220)	(0.0222)	(0.0225)	(0.0223)	(0.0226)
logavgsun			-0.00412	-0.00301	-0.00383	-0.00235	-0.00342	-0.00240	-0.00355	-0.00212
			(0.00903)	(0.00907)	(0.00911)	(0.00916)	(0.00919)	(0.00922)	(0.00930)	(0.00935)
domestic							0.000146	0.000121	0.000138	0.000136
							(0.000356)	(0.000363)	(0.000366)	(0.000371)
hmt							-0.000658	-0.000704	-0.000660	-0.000701
							(0.000426)	(0.000429)	(0.000427)	(0.000430)
foreign							0.000258	0.000280	0.000266	0.000266
							(0.000563)	(0.000570)	(0.000570)	(0.000576)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-1.820***	-1.864***	-1.869***	-1.920***	-1.886***	-1.957***	-1.842***	-1.887***	-1.835***	-1.900***
	(0.0627)	(0.0785)	(0.127)	(0.137)	(0.141)	(0.153)	(0.0912)	(0.104)	(0.111)	(0.124)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.896	0.897	0.904	0.904	0.904	0.905	0.901	0.902	0.901	0.902
Hausman test	-	-	0.0002	0.6125	0.0000	-	0.0000	0.0000	-	-

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

	Random Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10
logelec	-0.00494	-0.0148	0.000272	0.0154	-0.00434	0.0145	-0.00410	-0.0102	-0.00780	-0.00216
	(0.00804)	(0.0376)	(0.00725)	(0.0337)	(0.00696)	(0.0318)	(0.00832)	(0.0382)	(0.00787)	(0.0362)
logelec2		0.00149		-0.00188		-0.00241		0.00113		-0.000481
		(0.00509)		(0.00457)		(0.00431)		(0.00516)		(0.00490)
logcoalgas	-0.00870	0.0255	-0.00240	0.0384	-9.28e-05	0.0560**	-0.00242	0.0526*	-0.000301	0.0682***
	(0.00695)	(0.0273)	(0.00625)	(0.0246)	(0.00598)	(0.0235)	(0.00702)	(0.0276)	(0.00665)	(0.0264)
logcoalgas2		-0.00536		-0.00647*		-0.00891**		-0.00868**		-0.0108***
		(0.00413)		(0.00374)		(0.00356)		(0.00420)		(0.00401)
loglpg	0.0177**	0.0170	0.00971	-0.00315	0.00391	0.00262	0.0160*	0.00940	0.00812	0.00725
	(0.00860)	(0.0371)	(0.00769)	(0.0328)	(0.00741)	(0.0310)	(0.00868)	(0.0376)	(0.00834)	(0.0357)
loglpg2		-0.000153		0.00153		-0.000309		0.000559		-0.000480
		(0.00513)		(0.00453)		(0.00429)		(0.00520)		(0.00493)
loggdppercap			-0.0329***	-0.0354***	0.169***	0.180***	-0.0319***	-0.0356***	0.191***	0.200***
			(0.00875)	(0.00887)	(0.0403)	(0.0399)	(0.0105)	(0.0106)	(0.0459)	(0.0459)
loggdppercap2					-0.0270***	-0.0289***			-0.0301***	-0.0320***
					(0.00527)	(0.00525)			(0.00607)	(0.00609)
secondarygdp			0.000398	0.000440	0.000553	0.000608*				
			(0.000384)	(0.000383)	(0.000367)	(0.000364)				
servicegdp			-0.000706**	-0.000660*	-0.000423	-0.000345				
			(0.000358)	(0.000356)	(0.000346)	(0.000341)				
logmaxtemp			-0.0114	-0.00945	-0.0133	-0.00997	-0.00334	0.00104	-0.00648	-0.000354
			(0.0178)	(0.0179)	(0.0170)	(0.0170)	(0.0195)	(0.0195)	(0.0186)	(0.0186)
logmintemp			-0.00447	-0.00447	-0.00186	-0.00165	-0.00657	-0.00600	-0.00555	-0.00582
			(0.00952)	(0.00945)	(0.00908)	(0.00894)	(0.0108)	(0.0107)	(0.0102)	(0.0101)
logavgpre			-0.00835	-0.00918	-0.0110	-0.0123	-0.00718	-0.00787	-0.0109	-0.0117
			(0.00873)	(0.00867)	(0.00833)	(0.00821)	(0.00981)	(0.00972)	(0.00929)	(0.00924)
logavghumi			0.0566**	0.0548**	0.0533**	0.0490**	0.0182	0.0136	0.0197	0.0102
			(0.0227)	(0.0230)	(0.0218)	(0.0219)	(0.0246)	(0.0248)	(0.0236)	(0.0237)
logavgsun			-0.00341	-0.00303	0.000491	0.00176	0.00967	0.00960	0.0123	0.0137
			(0.01000)	(0.00995)	(0.00958)	(0.00947)	(0.0111)	(0.0110)	(0.0105)	(0.0105)
domestic							0.00104**	0.000977**	0.00113***	0.00106**
							(0.000455)	(0.000459)	(0.000429)	(0.000435)
hmt							5.83e-05	1.23e-05	-8.70e-05	-0.000149
							(0.000547)	(0.000543)	(0.000516)	(0.000515)
foreign							0.000887	0.001000	9.08e-05	0.000121
							(0.000704)	(0.000703)	(0.000688)	(0.000689)
Constant	-2.333***	-2.351***	-2.196***	-2.241***	-2.547***	-2.651***	-2.260***	-2.282***	-2.547***	-2.615***
	(0.0715)	(0.0973)	(0.145)	(0.157)	(0.154)	(0.167)	(0.110)	(0.127)	(0.118)	(0.136)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

Residential usage

Furthermore, our data enables us to decompose the energy consumption into residential usage and industrial usage to investigate the impact of the consumption of urban residents on air quality. In Table 6.4_E, results show that residential coal and gas consumption significantly positively affects the city's PM_{10} concentration, either in fixed or random effects (the impact is around 0.01 to 0.04).

In the fixed effects model, the output of HMT firms shows a significant negative effect on PM_{10} concentration (magnitude around 0.0008), which may indicate that firms from Hong Kong, Macau and Taiwan might be less pollution intensive than other firms. These firms tend to concentrate in high technology or light manufacturing industries. The random effects model shows that output from domestic firms and foreign firms have a significant positive impact on PM_{10} concentrations. However, the magnitudes of these effects are quite small (the magnitude for domestic firms is around 0.0009 and for foreign firms 0.0012).

Table 6.4_ E City Size by Energy Consumption- Decompose to Residential Energy Use

	Fixed Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10
logresielec	-0.000993	-0.0150	-0.000859	-0.000878	-0.0175	-0.0177	-0.00163	-0.00163	-0.0193	-0.0189
	(0.00660)	(0.0285)	(0.00659)	(0.00661)	(0.0290)	(0.0291)	(0.00681)	(0.00682)	(0.0294)	(0.0295)
logresielec2		0.00195			0.00231	0.00233			0.00244	0.00238
		(0.00386)			(0.00394)	(0.00395)			(0.00399)	(0.00400)
logcoalgasresi	0.0131**	0.00303	0.0109**	0.0109**	0.00381	0.00440	0.0120**	0.0120**	0.00162	0.000629
	(0.00547)	(0.0214)	(0.00551)	(0.00552)	(0.0212)	(0.0214)	(0.00553)	(0.00555)	(0.0214)	(0.0216)
logcoalgasresi2		0.00145			0.000994	0.000901			0.00146	0.00161
		(0.00323)			(0.00319)	(0.00323)			(0.00322)	(0.00325)
loglpgresi	-0.00553	0.0231	-0.00687	-0.00691	0.0218	0.0217	-0.00808	-0.00808	0.0324	0.0332
	(0.00679)	(0.0286)	(0.00671)	(0.00673)	(0.0284)	(0.0284)	(0.00701)	(0.00702)	(0.0293)	(0.0295)
loglpgresi2		-0.00415			-0.00415	-0.00415			-0.00590	-0.00601
		(0.00407)			(0.00403)	(0.00404)			(0.00419)	(0.00421)
loggdppercap			0.00175	0.0115	0.000350	0.00904	0.00204	-0.00819	0.000549	-0.0154
			(0.00857)	(0.0430)	(0.00887)	(0.0435)	(0.00881)	(0.0449)	(0.00910)	(0.0454)
loggdppercap2				-0.00140		-0.00125		0.00147		0.00230
				(0.00604)		(0.00612)		(0.00632)		(0.00640)
secondarygdp			0.000552	0.000557	0.000521	0.000524				
			(0.000342)	(0.000343)	(0.000345)	(0.000346)				
servicegdp			4.60e-06	4.87e-06	-8.20e-06	-8.31e-06				
			(0.000326)	(0.000326)	(0.000328)	(0.000328)				
logmaxtemp			-0.0191	-0.0192	-0.0189	-0.0190	-0.0234	-0.0234	-0.0235	-0.0235
			(0.0161)	(0.0162)	(0.0162)	(0.0162)	(0.0165)	(0.0166)	(0.0166)	(0.0166)
logmintemp			0.00570	0.00562	0.00597	0.00590	0.00537	0.00550	0.00598	0.00617
			(0.00839)	(0.00841)	(0.00843)	(0.00845)	(0.00843)	(0.00846)	(0.00846)	(0.00849)
logavgpre			-0.0121	-0.0123	-0.0129*	-0.0130*	-0.0131*	-0.0130*	-0.0139*	-0.0137*
			(0.00755)	(0.00759)	(0.00762)	(0.00765)	(0.00766)	(0.00771)	(0.00771)	(0.00774)
logavghumi			0.0647***	0.0647***	0.0655***	0.0655***	0.0663***	0.0665***	0.0683***	0.0686***
			(0.0217)	(0.0217)	(0.0218)	(0.0218)	(0.0222)	(0.0223)	(0.0223)	(0.0224)
logavgsun			-0.00323	-0.00300	-0.00308	-0.00288	-0.00258	-0.00288	-0.00265	-0.00310
			(0.00905)	(0.00912)	(0.00909)	(0.00915)	(0.00917)	(0.00928)	(0.00920)	(0.00930)
domestic							0.000136	0.000118	4.15e-05	1.14e-05
							(0.000354)	(0.000363)	(0.000360)	(0.000371)
hmt							-0.000822*	-0.000826*	-0.000830*	-0.000838*
							(0.000426)	(0.000427)	(0.000431)	(0.000432)
foreign							0.000369	0.000387	0.000464	0.000494
							(0.000573)	(0.000579)	(0.000581)	(0.000588)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-1.919***	-1.935***	-1.950***	-1.965***	-1.955***	-1.968***	-1.898***	-1.883***	-1.920***	-1.898***
	(0.0573)	(0.0755)	(0.128)	(0.142)	(0.141)	(0.154)	(0.0901)	(0.111)	(0.105)	(0.122)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.897	0.898	0.904	0.904	0.905	0.905	0.902	0.902	0.903	0.903
Hausman test	-	-	-	0.0000	0.0000	0.0000	0.0000	-	0.0000	-

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

	Random Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10	logpm_10
logresielec	-0.00507	-0.0382	0.000450	-0.000119	-0.0440	-0.0380	-0.00518	-0.00277	-0.0578	-0.0497
	(0.00841)	(0.0352)	(0.00753)	(0.00715)	(0.0321)	(0.0307)	(0.00851)	(0.00803)	(0.0360)	(0.0340)
logresielec2		0.00490			0.00627	0.00526			0.00746	0.00653
		(0.00475)			(0.00434)	(0.00417)			(0.00487)	(0.00461)
logcoalgasresi	0.0193***	-0.0434*	0.0139**	0.0139**	-0.0145	0.00575	0.0217***	0.0192***	-0.0106	0.00802
	(0.00695)	(0.0260)	(0.00628)	(0.00597)	(0.0235)	(0.0229)	(0.00689)	(0.00651)	(0.0263)	(0.0252)
logcoalgasresi2		0.00949**			0.00419	0.00106			0.00483	0.00159
		(0.00392)			(0.00355)	(0.00346)			(0.00397)	(0.00382)
loglpgresi	-0.0151*	0.0605*	-0.0104	-0.00984	0.0500	0.0375	-0.00635	-0.00768	0.0646*	0.0426
	(0.00857)	(0.0356)	(0.00766)	(0.00728)	(0.0318)	(0.0305)	(0.00878)	(0.00828)	(0.0365)	(0.0348)
loglpgresi2		-0.0106**			-0.00862*	-0.00683			-0.0102**	-0.00727
		(0.00504)			(0.00450)	(0.00432)			(0.00519)	(0.00494)
loggdppercap			-0.0320***	0.169***	-0.0331***	0.160***	-0.0311***	0.188***	-0.0328***	0.174***
			(0.00867)	(0.0386)	(0.00875)	(0.0389)	(0.0102)	(0.0433)	(0.0104)	(0.0440)
loggdppercap2				-0.0270***		-0.0260***		-0.0297***		-0.0281***
				(0.00506)		(0.00512)		(0.00573)		(0.00583)
secondarygdp			0.000500	0.000650*	0.000422	0.000579				
			(0.000381)	(0.000363)	(0.000378)	(0.000364)				
servicegdp			-0.000606*	-0.000307	-0.000607*	-0.000330				
			(0.000358)	(0.000345)	(0.000353)	(0.000343)				
logmaxtemp			-0.0162	-0.0157	-0.0165	-0.0151	-0.0101	-0.0112	-0.00979	-0.0106
			(0.0174)	(0.0166)	(0.0172)	(0.0165)	(0.0189)	(0.0179)	(0.0188)	(0.0179)
logmintemp			-0.00593	-0.00349	-0.00419	-0.00232	-0.00787	-0.00723	-0.00613	-0.00569
			(0.00936)	(0.00890)	(0.00922)	(0.00884)	(0.0104)	(0.00980)	(0.0103)	(0.00978)
logavgpre			-0.00875	-0.0122	-0.0116	-0.0139*	-0.00826	-0.0126	-0.0106	-0.0141
			(0.00852)	(0.00812)	(0.00844)	(0.00810)	(0.00943)	(0.00893)	(0.00939)	(0.00892)
logavghumi			0.0601***	0.0550**	0.0640***	0.0572***	0.0285	0.0277	0.0305	0.0301
			(0.0224)	(0.0214)	(0.0222)	(0.0214)	(0.0241)	(0.0231)	(0.0240)	(0.0230)
logavgsun			-0.00221	0.00159	-0.00198	0.00133	0.00894	0.0118	0.00945	0.0113
			(0.00991)	(0.00946)	(0.00981)	(0.00942)	(0.0108)	(0.0103)	(0.0108)	(0.0102)
domestic							0.00101**	0.00108***	0.000787*	0.000909**
							(0.000439)	(0.000414)	(0.000445)	(0.000422)
hmt							-0.000236	-0.000354	-0.000218	-0.000330
							(0.000532)	(0.000501)	(0.000533)	(0.000504)
foreign							0.00126*	0.000379	0.00129*	0.000468
							(0.000693)	(0.000679)	(0.000698)	(0.000687)
Constant	-2.304***	-2.299***	-2.198***	-2.585***	-2.174***	-2.552***	-2.253***	-2.567***	-2.232***	-2.528***
	(0.0736)	(0.0935)	(0.143)	(0.154)	(0.155)	(0.166)	(0.109)	(0.119)	(0.124)	(0.133)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

6.4.2 Results for SO_2

City Size by Population

Similarly we estimate SO_2 with the same regression specifications and find that firstly, in Table 6.5_A city size by population, there is a significant linear relationship between population size and SO_2 concentration. A city's population size has a significant negative effect on its SO_2 concentration levels (the magnitude is around 0.16-0.31) either in fixed or random effects[63]. This might indicate that the more population in a city, the less industrial activity within the city, thus, less SO_2 concentration.

City Size by Urban Area

Table 6.5_B shows the results when city size is represented by urban area (built-up area). For the role of urban built-up area, fixed effects show a linear positive significant correlation between urban area and SO_2 concentration. Random effects show a non-linear inverted-U shape between urban area and SO_2 concentrations.

City Size by Total GDP

There is a significant non-linear relationship between a city's total GDP size and the SO_2 concentration level since both fixed and random effects show significant negative estimated coefficients for the squared term of total GDP. Therefore there seems to be an inverted-U shape between a city's economic size and SO_2 concentration.

[63] One exception in random effects- regression (10), but the Hausman test shows that we reject the null hypothesis at 99% confidence level that random effects is preferred for the data.

City Size by Energy Consumption

LPG consumption is the main source of energy consumption in terms of its contribution to SO_2 emissions. Random effects results also show that LPG (liquefied petroleum gas) consumption has a significant linear positive effect on SO_2 concentration levels, a 1% increase in the consumption of LPG will increase SO_2 concentration levels by 2.8% to 4.29%.

Coal gas consumption consistently shows a significant linear negative correlation with the city SO_2 concentration level, which might be because of the lower SO_2 content of coal gas.

In addition, in random effects models, we find that the maximum temperature of a city shows a significant positive correlation with the SO_2 concentration level. This might indicate the important role of the use of air conditioners in city SO_2 concentration levels.

Residential Usage

When we decompose the energy consumption into residential usage and industrial usage, we find that residential electricity and LPG usage have significant linear positive effects on SO_2 concentration levels in the fixed effects regressions (4)(6)(7)(8)(10) (Hausman tests show these regressions prefer fixed effects). This might be because the supply of electricity in China is mainly based on the combustion of coal, which emits a large amount of $SO_{2,}$ and the LPG is the product from petroleum which is also a source of SO_2 emissions. The results indicate that if a city's residential electricity consumption increases by 1%, SO_2 concentration level will increase by 3.4%- 4.1% within that city. However, the residential coal gas usage has a significant negative effect on SO_2 concentration levels. This might be because the greater the use of coal gas for residential purposes, the lower the use of LPG and thus SO_2 concentrations are reduced.

The output from foreign firms shows a negative correlation with SO_2 concentration level, with estimated coefficients around -0.00227 to -0.00239, from all the estimates from fixed effects and one out of four estimates from random effects. This might show that foreign firms are cleaner in terms of emitting $SO_{2,}$ as if the share of output from foreign firms increase by 1%, city SO_2 concentration level might decrease by 0.002%.

Table 6.5_ A City Size by Population

	Fixed Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2
logpop	-0.310**	1.719	-0.290**	1.198	-0.292**	1.260	-0.286**	0.530	-0.286**	0.512
	(0.130)	(1.243)	(0.134)	(1.243)	(0.135)	(1.255)	(0.130)	(1.327)	(0.131)	(1.350)
logpop2		-0.125		-0.0922		-0.0962		-0.0504		-0.0492
		(0.0764)		(0.0765)		(0.0773)		(0.0815)		(0.0830)
loggdppercap			0.0222	0.0217	0.0445	0.0593	0.0196	0.0198	0.00134	0.0118
			(0.0189)	(0.0188)	(0.0953)	(0.0959)	(0.0194)	(0.0195)	(0.1000)	(0.102)
loggdppercap2					-0.00320	-0.00540			0.00262	0.00114
					(0.0134)	(0.0135)			(0.0141)	(0.0143)
secondarygdp			0.00236***	0.00237***	0.00237***	0.00239***				
			(0.000754)	(0.000753)	(0.000756)	(0.000756)				
servicegdp			0.00136*	0.00134*	0.00136*	0.00134*				
			(0.000730)	(0.000729)	(0.000731)	(0.000731)				
logmaxtemp			0.0265	0.0238	0.0264	0.0234	0.0138	0.0131	0.0138	0.0131
			(0.0359)	(0.0360)	(0.0360)	(0.0361)	(0.0369)	(0.0369)	(0.0369)	(0.0370)
logmintemp			-0.00955	-0.0105	-0.00971	-0.0108	-0.0185	-0.0189	-0.0183	-0.0188
			(0.0186)	(0.0186)	(0.0186)	(0.0186)	(0.0187)	(0.0188)	(0.0188)	(0.0188)
logavgpre			-0.0383**	-0.0353**	-0.0386**	-0.0356**	-0.0420**	-0.0403**	-0.0417**	-0.0403**
			(0.0168)	(0.0169)	(0.0168)	(0.0170)	(0.0170)	(0.0173)	(0.0171)	(0.0173)
logavghumi			0.0741	0.0683	0.0740	0.0679	0.0789	0.0745	0.0793	0.0748
			(0.0480)	(0.0481)	(0.0480)	(0.0482)	(0.0493)	(0.0499)	(0.0494)	(0.0501)
logavgsun			-0.00459	-0.00398	-0.00408	-0.00311	-0.00137	-0.000623	-0.00189	-0.000866
			(0.0199)	(0.0199)	(0.0200)	(0.0200)	(0.0203)	(0.0203)	(0.0205)	(0.0206)
domestic							-0.000583	-0.000554	-0.000616	-0.000569
							(0.000787)	(0.000789)	(0.000808)	(0.000813)
hmt							-0.000673	-0.000531	-0.000680	-0.000537
							(0.000932)	(0.000961)	(0.000935)	(0.000967)
foreign							-0.00145	-0.00144	-0.00142	-0.00143
							(0.00124)	(0.00124)	(0.00126)	(0.00126)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-0.248	-8.255	-1.000	-6.836	-1.016	-7.119	-0.471	-3.699	-0.452	-3.616
	(1.199)	(5.021)	(1.305)	(5.017)	(1.309)	(5.075)	(1.212)	(5.360)	(1.219)	(5.470)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.853	0.855	0.865	0.866	0.865	0.866	0.861	0.861	0.861	0.861
Hausman test	0.9404	-	-	-	0.0003	-	0.0000	0.0934	0.0000	0.0000

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

	Random Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2
logpop	-0.303***	0.616	-0.246***	0.642	-0.161*	1.488	-0.0995	1.547	-0.0451	2.212*
	(0.0898)	(1.243)	(0.0885)	(1.189)	(0.0908)	(1.177)	(0.0855)	(1.213)	(0.0847)	(1.187)
logpop2		-0.0568		-0.0518		-0.0988		-0.0965		-0.134*
		(0.0765)		(0.0729)		(0.0721)		(0.0738)		(0.0722)
loggdppercap			-0.0154	-0.0190	0.349***	0.372***	-0.0290	-0.0321	0.413***	0.444***
			(0.0193)	(0.0197)	(0.0882)	(0.0900)	(0.0212)	(0.0218)	(0.0928)	(0.0954)
loggdppercap2					-0.0493***	-0.0531***			-0.0599***	-0.0646***
					(0.0117)	(0.0120)			(0.0123)	(0.0126)
secondarygdp			0.00138*	0.00137	0.00170**	0.00172**				
			(0.000827)	(0.000838)	(0.000804)	(0.000811)				
servicegdp			-0.000172	-0.000181	0.000439	0.000461				
			(0.000773)	(0.000786)	(0.000762)	(0.000771)				
logmaxtemp			0.0546	0.0535	0.0519	0.0499	0.0404	0.0417	0.0402	0.0424
			(0.0381)	(0.0384)	(0.0370)	(0.0371)	(0.0396)	(0.0398)	(0.0381)	(0.0381)
logmintemp			-0.0380*	-0.0413**	-0.0346*	-0.0379*	-0.0459**	-0.0519**	-0.0461**	-0.0514**
			(0.0205)	(0.0208)	(0.0198)	(0.0201)	(0.0217)	(0.0223)	(0.0208)	(0.0212)
logavgpre			-0.0405**	-0.0383**	-0.0465**	-0.0434**	-0.0432**	-0.0398**	-0.0505***	-0.0467**
			(0.0186)	(0.0190)	(0.0181)	(0.0184)	(0.0197)	(0.0203)	(0.0189)	(0.0193)
logavghumi			0.0492	0.0390	0.0422	0.0290	0.0331	0.00717	0.0223	-0.00680
			(0.0487)	(0.0491)	(0.0473)	(0.0476)	(0.0507)	(0.0512)	(0.0489)	(0.0493)
logavgsun			0.00164	0.00523	0.0107	0.0149	0.0209	0.0297	0.0296	0.0386*
			(0.0216)	(0.0219)	(0.0211)	(0.0213)	(0.0227)	(0.0231)	(0.0219)	(0.0222)
domestic							0.000310	0.000426	0.000521	0.000686
							(0.000923)	(0.000954)	(0.000884)	(0.000907)
hmt							0.000910	0.00143	0.000772	0.00137
							(0.00110)	(0.00116)	(0.00106)	(0.00110)
foreign							-0.000183	-0.000227	-0.00177	-0.00187
							(0.00145)	(0.00149)	(0.00143)	(0.00146)
Constant	-0.686	-4.378	-1.254	-4.982	-2.653***	-9.475**	-2.161***	-9.074*	-3.247***	-12.65***
	(0.732)	(5.039)	(0.764)	(4.814)	(0.836)	(4.817)	(0.721)	(4.969)	(0.745)	(4.888)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

Table 6.5_ B City Size by Urban Area (Built-Up Area)

	Fixed Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2
logbuiltup	0.0397*	-0.0673	0.0430**	-0.0382	0.0442**	-0.0397	0.0445**	-0.149	0.0446**	-0.149
	(0.0202)	(0.120)	(0.0198)	(0.119)	(0.0200)	(0.120)	(0.0202)	(0.123)	(0.0205)	(0.124)
logbuiltup2		0.0162		0.0123		0.0127		0.0293		0.0295
		(0.0179)		(0.0178)		(0.0178)		(0.0184)		(0.0185)
loggdppercap			0.0216	0.0233	0.0623	0.0674	0.0193	0.0217	0.0236	0.0367
			(0.0189)	(0.0190)	(0.0960)	(0.0964)	(0.0194)	(0.0194)	(0.101)	(0.101)
loggdppercap2					-0.00584	-0.00633			-0.000626	-0.00214
					(0.0135)	(0.0135)			(0.0142)	(0.0142)
secondarygdp			0.00262***	0.00258***	0.00264***	0.00260***				
			(0.000752)	(0.000755)	(0.000754)	(0.000757)				
servicegdp			0.00188***	0.00187***	0.00188***	0.00188***				
			(0.000711)	(0.000712)	(0.000713)	(0.000713)				
logmaxtemp			0.0166	0.0175	0.0162	0.0172	0.00363	0.00433	0.00363	0.00434
			(0.0356)	(0.0357)	(0.0357)	(0.0358)	(0.0366)	(0.0365)	(0.0367)	(0.0366)
logmintemp			-0.0120	-0.0117	-0.0123	-0.0121	-0.0234	-0.0234	-0.0235	-0.0236
			(0.0185)	(0.0186)	(0.0186)	(0.0186)	(0.0187)	(0.0186)	(0.0188)	(0.0187)
logavgpre			-0.0424**	-0.0421**	-0.0430**	-0.0427**	-0.0460***	-0.0453***	-0.0460***	-0.0455***
			(0.0167)	(0.0167)	(0.0168)	(0.0168)	(0.0170)	(0.0170)	(0.0171)	(0.0171)
logavghumi			0.0856*	0.0857*	0.0856*	0.0857*	0.0922*	0.0951*	0.0921*	0.0949*
			(0.0476)	(0.0477)	(0.0477)	(0.0477)	(0.0490)	(0.0489)	(0.0491)	(0.0490)
logavgsun			-0.00236	-0.00281	-0.00144	-0.00182	-7.89e-06	-0.00200	0.000113	-0.00159
			(0.0198)	(0.0199)	(0.0200)	(0.0200)	(0.0202)	(0.0202)	(0.0204)	(0.0204)
domestic							-0.000619	-0.000621	-0.000611	-0.000594
							(0.000787)	(0.000784)	(0.000808)	(0.000806)
hmt							-0.000786	-0.00105	-0.000785	-0.00104
							(0.000937)	(0.000948)	(0.000939)	(0.000950)
foreign							-0.00148	-0.00172	-0.00149	-0.00174
							(0.00124)	(0.00125)	(0.00126)	(0.00126)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-3.215***	-3.037***	-3.887***	-3.760***	-3.950***	-3.825***	-3.183***	-2.870***	-3.190***	-2.890***
	(0.0966)	(0.220)	(0.271)	(0.328)	(0.309)	(0.356)	(0.182)	(0.268)	(0.237)	(0.301)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.852	0.853	0.865	0.865	0.865	0.865	0.861	0.862	0.861	0.862
Hausman test	-	-	-	-	0.0000	0.0000	0.0017	0.1359	0.0000	0.0000

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

	Random Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2
logbuiltup	0.000363	0.287**	0.00973	0.252**	0.0357	0.123	0.00639	0.283**	0.0335	0.122
	(0.0230)	(0.130)	(0.0222)	(0.126)	(0.0217)	(0.124)	(0.0234)	(0.131)	(0.0229)	(0.130)
logbuiltup2		-0.0425**		-0.0360*		-0.0131		-0.0410**		-0.0133
		(0.0190)		(0.0184)		(0.0183)		(0.0191)		(0.0192)
loggdppercap			-0.0286	-0.0269	0.411***	0.396***	-0.0318	-0.0284	0.448***	0.429***
			(0.0192)	(0.0192)	(0.0879)	(0.0909)	(0.0215)	(0.0214)	(0.0942)	(0.0976)
loggdppercap2					-0.0593***	-0.0571***			-0.0654***	-0.0626***
					(0.0116)	(0.0120)			(0.0125)	(0.0131)
secondarygdp			0.00148*	0.00161*	0.00188**	0.00191**				
			(0.000838)	(0.000838)	(0.000804)	(0.000807)				
servicegdp			8.46e-06	0.000152	0.000706	0.000731				
			(0.000783)	(0.000783)	(0.000758)	(0.000761)				
logmaxtemp			0.0415	0.0401	0.0431	0.0428	0.0382	0.0360	0.0397	0.0386
			(0.0383)	(0.0381)	(0.0367)	(0.0368)	(0.0397)	(0.0394)	(0.0380)	(0.0380)
logmintemp			-0.0418**	-0.0408**	-0.0371*	-0.0371*	-0.0485**	-0.0465**	-0.0492**	-0.0480**
			(0.0208)	(0.0207)	(0.0198)	(0.0199)	(0.0219)	(0.0218)	(0.0209)	(0.0208)
logavgpre			-0.0401**	-0.0419**	-0.0480***	-0.0482***	-0.0424**	-0.0449**	-0.0509***	-0.0515***
			(0.0189)	(0.0188)	(0.0181)	(0.0181)	(0.0199)	(0.0198)	(0.0190)	(0.0190)
logavghumi			0.0582	0.0575	0.0479	0.0470	0.0279	0.0305	0.0167	0.0198
			(0.0492)	(0.0489)	(0.0473)	(0.0473)	(0.0507)	(0.0504)	(0.0487)	(0.0488)
logavgsun			0.00784	0.00827	0.0151	0.0152	0.0252	0.0250	0.0321	0.0313
			(0.0218)	(0.0217)	(0.0209)	(0.0210)	(0.0227)	(0.0225)	(0.0217)	(0.0217)
domestic							0.000412	0.000226	0.000622	0.000557
							(0.000930)	(0.000926)	(0.000885)	(0.000888)
hmt							0.00105	0.00126	0.000731	0.000797
							(0.00111)	(0.00111)	(0.00106)	(0.00106)
foreign							1.90e-05	0.000118	-0.00179	-0.00164
							(0.00144)	(0.00143)	(0.00142)	(0.00142)
Constant	-3.141***	-3.547***	-3.255***	-3.636***	-4.220***	-4.321***	-2.975***	-3.383***	-3.778***	-3.882***
	(0.135)	(0.225)	(0.308)	(0.364)	(0.349)	(0.379)	(0.217)	(0.286)	(0.257)	(0.293)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

Table 6.5_ C City Size by Total GDP

Fixed Effects										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2
loggdp	-0.243*	0.0103	-0.385***	-0.126	-0.386***	-0.128	-0.286**	0.0753	-0.298**	0.0633
	(0.127)	(0.134)	(0.148)	(0.153)	(0.148)	(0.153)	(0.137)	(0.167)	(0.139)	(0.168)
loggdp2		-0.0651***		-0.0632***		-0.0641***		-0.0658***		-0.0658***
		(0.0139)		(0.0139)		(0.0140)		(0.0181)		(0.0182)
loggdppercap			0.0280	0.00886	0.0229	-0.0413	0.0234	0.00958	-0.0337	-0.0499
			(0.0189)	(0.0187)	(0.0948)	(0.0924)	(0.0195)	(0.0194)	(0.101)	(0.0986)
loggdppercap2					0.000725	0.00717			0.00822	0.00855
					(0.0133)	(0.0129)			(0.0142)	(0.0139)
secondarygdp			0.00272***	0.00278***	0.00271***	0.00276***				
			(0.000751)	(0.000723)	(0.000753)	(0.000725)				
servicegdp			0.00136*	0.00141**	0.00135*	0.00141**				
			(0.000721)	(0.000694)	(0.000722)	(0.000695)				
logmaxtemp			0.0243	0.00986	0.0243	0.0102	0.00676	0.00475	0.00679	0.00478
			(0.0356)	(0.0345)	(0.0357)	(0.0345)	(0.0367)	(0.0358)	(0.0367)	(0.0359)
logmintemp			-0.0129	-0.0154	-0.0128	-0.0151	-0.0219	-0.0233	-0.0212	-0.0226
			(0.0185)	(0.0178)	(0.0185)	(0.0178)	(0.0187)	(0.0182)	(0.0187)	(0.0183)
logavgpre			-0.0379**	-0.0351**	-0.0379**	-0.0343**	-0.0429**	-0.0398**	-0.0419**	-0.0388**
			(0.0167)	(0.0161)	(0.0168)	(0.0162)	(0.0170)	(0.0167)	(0.0171)	(0.0168)
logavghumi			0.0541	0.0593	0.0540	0.0589	0.0692	0.0629	0.0694	0.0631
			(0.0490)	(0.0472)	(0.0491)	(0.0473)	(0.0501)	(0.0490)	(0.0502)	(0.0490)
logavgsun			0.00404	0.00279	0.00392	0.00167	0.00602	0.00706	0.00454	0.00552
			(0.0199)	(0.0191)	(0.0200)	(0.0193)	(0.0203)	(0.0199)	(0.0205)	(0.0200)
domestic							-0.000750	0.000344	-0.000860	0.000231
							(0.000790)	(0.000829)	(0.000814)	(0.000850)
hmt							-0.000926	0.000260	-0.000967	0.000219
							(0.000949)	(0.000982)	(0.000953)	(0.000986)
foreign							-0.00152	-0.000448	-0.00142	-0.000344
							(0.00124)	(0.00125)	(0.00126)	(0.00126)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-2.208***	-2.067***	-2.265***	-2.058***	-2.256***	-1.960***	-2.029***	-2.267***	-1.900***	-2.133***
	(0.480)	(0.463)	(0.630)	(0.609)	(0.656)	(0.634)	(0.544)	(0.535)	(0.589)	(0.578)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.852	0.864	0.866	0.876	0.866	0.877	0.861	0.868	0.861	0.868
Hausman test	0.9011	-	-	-	-	-	-	-	-	0.0000

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

Random Effects										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2
loggdp	-0.228***	0.163**	-0.231***	0.155*	-0.211***	0.158*	-0.263***	0.238*	-0.236***	0.263**
	(0.0271)	(0.0785)	(0.0326)	(0.0865)	(0.0416)	(0.0877)	(0.0371)	(0.126)	(0.0478)	(0.129)
loggdp2		-0.0734***		-0.0686***		-0.0682***		-0.0790***		-0.0789***
		(0.0138)		(0.0144)		(0.0145)		(0.0189)		(0.0189)
loggdppercap			0.0349*	0.0105	0.112	0.0385	0.0203	0.00223	0.121	0.0956
			(0.0196)	(0.0196)	(0.0983)	(0.0964)	(0.0207)	(0.0205)	(0.106)	(0.103)
loggdppercap2					-0.0111	-0.00393			-0.0143	-0.0133
					(0.0138)	(0.0135)			(0.0149)	(0.0145)
secondarygdp			0.00184**	0.00195***	0.00187**	0.00196***				
			(0.000770)	(0.000746)	(0.000770)	(0.000745)				
servicegdp			0.000882	0.00110	0.000932	0.00112				
			(0.000727)	(0.000706)	(0.000728)	(0.000706)				
logmaxtemp			0.0576	0.0381	0.0564	0.0373	0.0402	0.0353	0.0399	0.0350
			(0.0354)	(0.0344)	(0.0354)	(0.0344)	(0.0365)	(0.0355)	(0.0365)	(0.0355)
logmintemp			-0.0296	-0.0320*	-0.0294	-0.0316*	-0.0412**	-0.0423**	-0.0413**	-0.0424**
			(0.0191)	(0.0185)	(0.0190)	(0.0184)	(0.0199)	(0.0193)	(0.0198)	(0.0192)
logavgpre			-0.0438**	-0.0404**	-0.0449***	-0.0409**	-0.0461**	-0.0429**	-0.0476***	-0.0443**
			(0.0173)	(0.0168)	(0.0174)	(0.0168)	(0.0181)	(0.0176)	(0.0181)	(0.0176)
logavghumi			0.0247	0.0373	0.0260	0.0386	0.0215	0.0190	0.0212	0.0187
			(0.0457)	(0.0442)	(0.0457)	(0.0443)	(0.0470)	(0.0457)	(0.0470)	(0.0457)
logavgsun			0.00615	0.00470	0.00774	0.00508	0.0153	0.0164	0.0176	0.0186
			(0.0200)	(0.0194)	(0.0201)	(0.0195)	(0.0208)	(0.0202)	(0.0209)	(0.0203)
o.logavgpre			-	-	-	-	-	-	-	-
domestic							-0.00101	0.000420	-0.000821	0.000592
							(0.000865)	(0.000906)	(0.000885)	(0.000923)
hmt							-0.000125	0.00127	-6.43e-05	0.00132
							(0.00102)	(0.00104)	(0.00102)	(0.00104)
foreign							-0.00233*	-0.000957	-0.00245*	-0.00108
							(0.00136)	(0.00135)	(0.00136)	(0.00136)
Constant	-2.514***	-2.954***	-3.048***	-3.407***	-3.223***	-3.468***	-2.353***	-3.018***	-2.574***	-3.222***
	(0.118)	(0.144)	(0.274)	(0.275)	(0.351)	(0.342)	(0.204)	(0.255)	(0.308)	(0.338)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

Table 6.5_ D City Size by Energy Consumption

	Fixed Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2
logelec	-0.00717	-0.0747	-0.00166	-0.0358	-0.00176	-0.0362	0.000901	-0.0440	0.00111	-0.0439
	(0.0138)	(0.0649)	(0.0140)	(0.0647)	(0.0141)	(0.0648)	(0.0143)	(0.0662)	(0.0144)	(0.0663)
logelec2		0.00927		0.00469		0.00471		0.00611		0.00611
		(0.00878)		(0.00878)		(0.00879)		(0.00896)		(0.00898)
logcoalgas	-0.0343***	-0.0169	-0.0338***	-0.0179	-0.0338***	-0.0154	-0.0357***	-0.0219	-0.0357***	-0.0227
	(0.0121)	(0.0471)	(0.0120)	(0.0472)	(0.0121)	(0.0481)	(0.0122)	(0.0479)	(0.0122)	(0.0487)
logcoalgas2		-0.00262		-0.00256		-0.00296		-0.00222		-0.00209
		(0.00712)		(0.00715)		(0.00730)		(0.00725)		(0.00740)
loglpg	0.0133	0.0751	0.0126	0.0682	0.0125	0.0690	0.00918	0.0856	0.00926	0.0855
	(0.0150)	(0.0638)	(0.0148)	(0.0627)	(0.0148)	(0.0629)	(0.0153)	(0.0645)	(0.0153)	(0.0647)
loglpg2		-0.00885		-0.00799		-0.00815		-0.0109		-0.0109
		(0.00884)		(0.00869)		(0.00873)		(0.00895)		(0.00897)
loggdppercap			0.0211	0.0191	0.0326	0.0461	0.0195	0.0178	-0.000883	0.00810
			(0.0191)	(0.0196)	(0.0954)	(0.0973)	(0.0196)	(0.0201)	(0.100)	(0.102)
loggdppercap2					-0.00164	-0.00389			0.00292	0.00140
					(0.0134)	(0.0137)			(0.0141)	(0.0144)
secondarygdp			0.00236***	0.00230***	0.00237***	0.00231***				
			(0.000751)	(0.000758)	(0.000754)	(0.000761)				
servicegdp			0.00167**	0.00167**	0.00167**	0.00167**				
			(0.000709)	(0.000712)	(0.000710)	(0.000714)				
logmaxtemp			0.0302	0.0344	0.0300	0.0343	0.0206	0.0251	0.0206	0.0250
			(0.0361)	(0.0366)	(0.0362)	(0.0366)	(0.0372)	(0.0375)	(0.0372)	(0.0376)
logmintemp			-0.00892	-0.00809	-0.00899	-0.00828	-0.0192	-0.0173	-0.0190	-0.0172
			(0.0186)	(0.0187)	(0.0186)	(0.0188)	(0.0187)	(0.0189)	(0.0188)	(0.0189)
logavgpre			-0.0423**	-0.0427**	-0.0424**	-0.0431**	-0.0461***	-0.0462***	-0.0458***	-0.0460***
			(0.0170)	(0.0170)	(0.0170)	(0.0171)	(0.0172)	(0.0173)	(0.0173)	(0.0174)
logavghumi			0.0824*	0.0775	0.0824*	0.0771	0.0868*	0.0831*	0.0871*	0.0834*
			(0.0476)	(0.0484)	(0.0477)	(0.0485)	(0.0490)	(0.0497)	(0.0491)	(0.0499)
logavgsun			-0.00534	-0.00512	-0.00507	-0.00438	-0.00304	-0.00346	-0.00364	-0.00377
			(0.0199)	(0.0200)	(0.0200)	(0.0202)	(0.0202)	(0.0203)	(0.0205)	(0.0206)
domestic							-0.000553	-0.000762	-0.000590	-0.000779
							(0.000784)	(0.000800)	(0.000806)	(0.000819)
hmt							-0.000458	-0.000407	-0.000467	-0.000411
							(0.000938)	(0.000946)	(0.000940)	(0.000948)
foreign							-0.00137	-0.00112	-0.00134	-0.00110
							(0.00124)	(0.00126)	(0.00126)	(0.00127)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-2.971***	-2.998***	-3.645***	-3.694***	-3.662***	-3.736***	-3.011***	-3.086***	-2.982***	-3.072***
	(0.139)	(0.174)	(0.280)	(0.302)	(0.310)	(0.337)	(0.201)	(0.228)	(0.244)	(0.273)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.855	0.856	0.867	0.867	0.867	0.868	0.863	0.864	0.863	0.864
Hausman test	-	-	-	-	0.0000	0.0000	-	0.0000	0.0000	0.4008

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

	Random Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2
logelec	0.00162	-0.0807	0.0135	-0.0301	0.00495	-0.0319	0.00897	-0.0687	0.00209	-0.0523
	(0.0159)	(0.0747)	(0.0154)	(0.0723)	(0.0150)	(0.0694)	(0.0165)	(0.0770)	(0.0159)	(0.0747)
logelec2		0.0114		0.00617		0.00522		0.0110		0.00780
		(0.0101)		(0.00980)		(0.00942)		(0.0104)		(0.0101)
logcoalgas	-0.0479***	-0.0335	-0.0434***	-0.00837	-0.0391***	0.0233	-0.0445***	0.00920	-0.0396***	0.0408
	(0.0138)	(0.0542)	(0.0133)	(0.0528)	(0.0129)	(0.0513)	(0.0140)	(0.0559)	(0.0135)	(0.0545)
logcoalgas2		-0.00206		-0.00551		-0.00988		-0.00818		-0.0124
		(0.00820)		(0.00802)		(0.00777)		(0.00849)		(0.00827)
loglpg	0.0415**	0.0742	0.0346**	0.0568	0.0237	0.0679	0.0429**	0.0763	0.0280*	0.0719
	(0.0170)	(0.0735)	(0.0164)	(0.0704)	(0.0160)	(0.0676)	(0.0172)	(0.0760)	(0.0169)	(0.0736)
loglpg2		-0.00464		-0.00329		-0.00673		-0.00499		-0.00683
		(0.0102)		(0.00972)		(0.00937)		(0.0105)		(0.0102)
loggdppercap			-0.0270	-0.0304	0.348***	0.361***	-0.0285	-0.0335	0.379***	0.389***
			(0.0186)	(0.0190)	(0.0869)	(0.0872)	(0.0209)	(0.0214)	(0.0930)	(0.0948)
loggdppercap2					-0.0501***	-0.0526***			-0.0552***	-0.0574***
					(0.0114)	(0.0115)			(0.0123)	(0.0126)
secondarygdp			0.00136*	0.00133	0.00166**	0.00165**				
			(0.000818)	(0.000822)	(0.000793)	(0.000794)				
servicegdp			6.44e-05	9.63e-05	0.000595	0.000676				
			(0.000764)	(0.000765)	(0.000746)	(0.000746)				
logmaxtemp			0.0625*	0.0674*	0.0603	0.0679*	0.0600	0.0683*	0.0570	0.0674*
			(0.0380)	(0.0384)	(0.0368)	(0.0371)	(0.0393)	(0.0398)	(0.0380)	(0.0385)
logmintemp			-0.0377*	-0.0361*	-0.0328*	-0.0309	-0.0436**	-0.0421*	-0.0430**	-0.0433**
			(0.0203)	(0.0203)	(0.0196)	(0.0195)	(0.0214)	(0.0216)	(0.0206)	(0.0209)
logavgpre			-0.0449**	-0.0456**	-0.0497***	-0.0511***	-0.0454**	-0.0452**	-0.0514***	-0.0516***
			(0.0186)	(0.0186)	(0.0180)	(0.0179)	(0.0195)	(0.0197)	(0.0189)	(0.0191)
logavghumi			0.0478	0.0437	0.0402	0.0312	0.0146	0.00197	0.00885	-0.0109
			(0.0484)	(0.0493)	(0.0470)	(0.0478)	(0.0501)	(0.0508)	(0.0485)	(0.0491)
logavgsun			0.00378	0.00314	0.0110	0.0119	0.0213	0.0216	0.0279	0.0307
			(0.0213)	(0.0214)	(0.0207)	(0.0207)	(0.0222)	(0.0223)	(0.0214)	(0.0216)
domestic							0.000338	0.000105	0.000515	0.000254
							(0.000902)	(0.000926)	(0.000868)	(0.000898)
hmt							0.00121	0.00128	0.000964	0.00102
							(0.00109)	(0.00110)	(0.00104)	(0.00106)
foreign							-1.84e-05	0.000154	-0.00158	-0.00153
							(0.00140)	(0.00142)	(0.00140)	(0.00142)
Constant	-3.129***	-3.072***	-3.272***	-3.268***	-3.928***	-4.016***	-3.044***	-3.017***	-3.563***	-3.612***
	(0.144)	(0.195)	(0.308)	(0.337)	(0.333)	(0.364)	(0.220)	(0.258)	(0.242)	(0.282)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

Table 6.5_ E City Size by Energy Consumption- Decomposed to Residential Energy Use

	Fixed Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2
logresielec	0.0342**	0.0630	0.0345**	0.0345**	0.0871	0.0873	0.0411***	0.0411***	0.0719	0.0726
	(0.0145)	(0.0627)	(0.0144)	(0.0144)	(0.0634)	(0.0635)	(0.0149)	(0.0149)	(0.0644)	(0.0645)
logresielec2		-0.00385			-0.00721	-0.00722			-0.00408	-0.00418
		(0.00849)			(0.00862)	(0.00864)			(0.00873)	(0.00876)
logcoalgasresi	-0.0229*	-0.0407	-0.0219*	-0.0219*	-0.0421	-0.0425	-0.0269**	-0.0269**	-0.0534	-0.0553
	(0.0120)	(0.0472)	(0.0120)	(0.0121)	(0.0463)	(0.0468)	(0.0121)	(0.0121)	(0.0468)	(0.0473)
logcoalgasresi2		0.00289			0.00322	0.00329			0.00421	0.00450
		(0.00711)			(0.00698)	(0.00706)			(0.00705)	(0.00713)
loglpgresi	0.0312**	0.00399	0.0283*	0.0283*	-0.00270	-0.00263	0.0246	0.0246	0.000752	0.00224
	(0.0149)	(0.0631)	(0.0147)	(0.0147)	(0.0620)	(0.0621)	(0.0153)	(0.0153)	(0.0642)	(0.0645)
loglpgresi2		0.00412			0.00464	0.00463			0.00366	0.00345
		(0.00897)			(0.00880)	(0.00882)			(0.00917)	(0.00921)
loggdppercap			0.0171	0.0186	0.0218	0.0157	0.0131	-0.0113	0.0162	-0.0148
			(0.0187)	(0.0939)	(0.0194)	(0.0950)	(0.0192)	(0.0980)	(0.0199)	(0.0994)
loggdppercap2				-0.000213		0.000872		0.00349		0.00446
				(0.0132)		(0.0134)		(0.0138)		(0.0140)
secondarygdp			0.00222***	0.00222***	0.00228***	0.00228***				
			(0.000747)	(0.000749)	(0.000754)	(0.000757)				
servicegdp			0.00135*	0.00135*	0.00140*	0.00140*				
			(0.000711)	(0.000712)	(0.000716)	(0.000717)				
logmaxtemp			0.0254	0.0254	0.0248	0.0248	0.0144	0.0144	0.0140	0.0139
			(0.0352)	(0.0353)	(0.0354)	(0.0355)	(0.0361)	(0.0361)	(0.0362)	(0.0363)
logmintemp			-0.0109	-0.0109	-0.0116	-0.0116	-0.0206	-0.0203	-0.0214	-0.0210
			(0.0183)	(0.0184)	(0.0184)	(0.0185)	(0.0184)	(0.0185)	(0.0185)	(0.0186)
logavgpre			-0.0439***	-0.0439***	-0.0435***	-0.0434***	-0.0483***	-0.0479***	-0.0485***	-0.0481***
			(0.0165)	(0.0166)	(0.0166)	(0.0167)	(0.0167)	(0.0168)	(0.0169)	(0.0170)
logavghumi			0.0773	0.0773	0.0761	0.0760	0.0803*	0.0807*	0.0791	0.0797
			(0.0474)	(0.0475)	(0.0476)	(0.0477)	(0.0485)	(0.0487)	(0.0489)	(0.0490)
logavgsun			-0.000836	-0.000801	-0.000591	-0.000726	0.00288	0.00218	0.00350	0.00263
			(0.0198)	(0.0199)	(0.0199)	(0.0200)	(0.0200)	(0.0202)	(0.0201)	(0.0204)
domestic							-0.000536	-0.000580	-0.000455	-0.000513
							(0.000772)	(0.000793)	(0.000789)	(0.000811)
hmt							-0.000251	-0.000261	-0.000248	-0.000264
							(0.000930)	(0.000932)	(0.000944)	(0.000946)
foreign							-0.00231*	-0.00227*	-0.00239*	-0.00233*
							(0.00125)	(0.00126)	(0.00127)	(0.00129)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-3.267***	-3.255***	-3.781***	-3.783***	-3.825***	-3.816***	-3.205***	-3.170***	-3.204***	-3.161***
	(0.126)	(0.166)	(0.279)	(0.310)	(0.308)	(0.337)	(0.197)	(0.242)	(0.230)	(0.267)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.858	0.859	0.870	0.870	0.871	0.871	0.867	0.867	0.868	0.868
Hausman test	-	-	-	0.0000	-	0.0000	0.0001	0.0000	-	0.0000

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

	Random Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2	logso_2
logresielec	0.0200	0.0794	0.0300*	0.0290*	0.0881	0.102	0.0243	0.0300*	0.0441	0.0626
	(0.0173)	(0.0741)	(0.0165)	(0.0158)	(0.0717)	(0.0687)	(0.0177)	(0.0168)	(0.0763)	(0.0743)
logresielec2		-0.00769			-0.00774	-0.00996			-0.00230	-0.00431
		(0.0100)			(0.00971)	(0.00932)			(0.0103)	(0.0101)
logcoalgasresi	-0.0180	-0.122**	-0.0229*	-0.0229*	-0.0916*	-0.0508	-0.0164	-0.0211	-0.0873	-0.0380
	(0.0143)	(0.0547)	(0.0138)	(0.0132)	(0.0525)	(0.0511)	(0.0143)	(0.0136)	(0.0557)	(0.0549)
logcoalgasresi2		0.0161*			0.0106	0.00427			0.0111	0.00306
		(0.00824)			(0.00794)	(0.00773)			(0.00841)	(0.00832)
loglpgresi	0.0204	0.0605	0.0253	0.0263	0.0458	0.0188	0.0339*	0.0305*	0.0579	0.00874
	(0.0176)	(0.0749)	(0.0168)	(0.0161)	(0.0710)	(0.0683)	(0.0183)	(0.0173)	(0.0773)	(0.0759)
loglpgresi2		-0.00516			-0.00267	0.00116			-0.00317	0.00327
		(0.0106)			(0.0101)	(0.00967)			(0.0110)	(0.0108)
loggdppercap			-0.0316*	0.377***	-0.0259	0.379***	-0.0368*	0.426***	-0.0333	0.430***
			(0.0190)	(0.0853)	(0.0195)	(0.0871)	(0.0212)	(0.0906)	(0.0220)	(0.0961)
loggdppercap2				-0.0547***		-0.0544***		-0.0627***		-0.0632***
				(0.0112)		(0.0114)		(0.0120)		(0.0127)
secondarygdp			0.00120	0.00152*	0.00128	0.00162**				
			(0.000833)	(0.000800)	(0.000841)	(0.000810)				
servicegdp			-0.000364	0.000248	-0.000269	0.000313				
			(0.000784)	(0.000761)	(0.000789)	(0.000766)				
logmaxtemp			0.0465	0.0486	0.0435	0.0464	0.0431	0.0443	0.0406	0.0451
			(0.0379)	(0.0364)	(0.0381)	(0.0366)	(0.0394)	(0.0376)	(0.0396)	(0.0382)
logmintemp			-0.0407**	-0.0355*	-0.0412**	-0.0365*	-0.0458**	-0.0452**	-0.0476**	-0.0512**
			(0.0205)	(0.0197)	(0.0206)	(0.0198)	(0.0216)	(0.0205)	(0.0219)	(0.0213)
logavgpre			-0.0416**	-0.0485***	-0.0420**	-0.0474***	-0.0435**	-0.0521***	-0.0444**	-0.0506***
			(0.0187)	(0.0179)	(0.0188)	(0.0181)	(0.0196)	(0.0187)	(0.0199)	(0.0194)
logavghumi			0.0568	0.0460	0.0560	0.0455	0.0257	0.0172	0.0212	-0.000446
			(0.0487)	(0.0470)	(0.0488)	(0.0471)	(0.0505)	(0.0484)	(0.0505)	(0.0486)
logavgsun			0.00775	0.0151	0.0105	0.0161	0.0261	0.0329	0.0304	0.0393*
			(0.0216)	(0.0208)	(0.0218)	(0.0210)	(0.0225)	(0.0215)	(0.0228)	(0.0221)
domestic							0.000515	0.000686	0.000458	0.000728
							(0.000913)	(0.000866)	(0.000945)	(0.000921)
hmt							0.00143	0.00118	0.00145	0.00130
							(0.00111)	(0.00105)	(0.00113)	(0.00110)
foreign							-0.000270	-0.00219	-0.000425	-0.00260*
							(0.00144)	(0.00142)	(0.00148)	(0.00150)
Constant	-3.248***	-3.304***	-3.274***	-4.069***	-3.373***	-4.175***	-3.148***	-3.811***	-3.152***	-3.796***
	(0.148)	(0.189)	(0.310)	(0.339)	(0.340)	(0.368)	(0.227)	(0.250)	(0.263)	(0.288)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

Composition Effect and City Climate Conditions

In all of the regression specifications, the composition of secondary industry over total GDP plays a significant positive role in a city's SO_2 concentrations. Specifically, both fixed and random effects show that an increase in the share of secondary industry will increase SO_2 concentrations in a city. In addition, fixed effects models also show that the service industry has a positive significant effect on SO_2 concentrations as well.

Results also show the important role of precipitation on SO_2 concentrations. Both fixed and random effects specification show a significant negative relationship between a city's precipitation and its SO_2 concentrations.

6.4.3 Results for NO_2

City Size by Population

From Table 6.6_A, we can find that both fixed and random effects models show an inverted-U shape between a city's population size and its NO_2 concentration level. The secondary industry or the outputs from domestic/HMT/foreign firms do not show significant correlations with $NO_{2.}$

It also shows that NO_2 concentrations tend to be negatively correlated with the annual maximum temperature in a city and positively correlated with the annual average sunshine hours.

City Size by Urban Area

Table 6.6_B shows that urban area (built-up area) has a significant linear correlation with a city's NO_2 concentration. When the built-up area increases by 1% the NO_2 concentration will increase by 4.42% to 5% (Hausman test shows that regression (1) to (7) cannot reject the null hypothesis that random effects are preferred to fit the data, therefore, we use the magnitudes from the random effects results).

City Size by Total GDP

Table 6.6_C shows that the total GDP of a city has a non-linear correlation with NO_2 concentrations. Specifically, both fixed and random effects show an inverted-U shape between total GDP of a city and its NO_2 concentrations. Output of HMT firms shows a positive relationship with the NO_2 concentrations in random effects specifications.

City Size by Energy Consumption

Table 6.6_ D indicates that only the consumption of LPG shows a significant impact on a city's NO_2 concentration levels. Specifically, there seems to be an inverted-u shape between the LPG consumption and NO_2 concentrations, since both the fixed and random effects show significant negative estimated coefficients for the squared term of the consumption of LPG.

Residential energy consumption

For the impact of residential energy usage on NO_2 concentration, we find that from Table 6.6_E, only electricity usage for residents in urban areas positively affects the NO_2 concentrations significantly.

Table 6.6_A City Size by Population

	Fixed Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2
logpop	-0.0783	1.757**	-0.0688	1.861**	-0.0739	1.998**	-0.0556	2.008**	-0.0576	2.179**
	(0.0844)	(0.806)	(0.0898)	(0.824)	(0.0900)	(0.829)	(0.0859)	(0.864)	(0.0860)	(0.876)
logpop2		-0.113**		-0.120**		-0.128**		-0.127**		-0.138**
		(0.0495)		(0.0508)		(0.0511)		(0.0531)		(0.0539)
loggdppercap			0.0159	0.0153	0.0780	0.0978	0.0164	0.0168	0.0614	0.0908
			(0.0126)	(0.0125)	(0.0636)	(0.0634)	(0.0128)	(0.0127)	(0.0657)	(0.0660)
loggdppercap2					-0.00891	-0.0118			-0.00646	-0.0106
					(0.00894)	(0.00892)			(0.00925)	(0.00929)
secondarygdp			0.000276	0.000284	0.000299	0.000316				
			(0.000504)	(0.000500)	(0.000505)	(0.000500)				
servicegdp			6.08e-05	3.31e-05	5.49e-05	2.32e-05				
			(0.000488)	(0.000484)	(0.000488)	(0.000483)				
logmaxtemp			-0.0374	-0.0409*	-0.0377	-0.0416*	-0.0402*	-0.0420*	-0.0401*	-0.0420*
			(0.0240)	(0.0239)	(0.0240)	(0.0238)	(0.0243)	(0.0240)	(0.0243)	(0.0240)
logmintemp			0.00880	0.00757	0.00836	0.00689	0.00922	0.00820	0.00870	0.00725
			(0.0124)	(0.0123)	(0.0124)	(0.0123)	(0.0123)	(0.0122)	(0.0124)	(0.0122)
logavgpre			-0.0138	-0.00996	-0.0146	-0.0107	-0.0140	-0.00991	-0.0147	-0.0107
			(0.0112)	(0.0112)	(0.0112)	(0.0112)	(0.0112)	(0.0112)	(0.0113)	(0.0113)
logavghumi			0.0282	0.0207	0.0280	0.0200	0.0302	0.0190	0.0292	0.0165
			(0.0321)	(0.0319)	(0.0321)	(0.0319)	(0.0324)	(0.0325)	(0.0325)	(0.0325)
logavgsun			0.0231*	0.0238*	0.0245*	0.0258*	0.0239*	0.0258*	0.0252*	0.0281**
			(0.0133)	(0.0132)	(0.0134)	(0.0132)	(0.0133)	(0.0132)	(0.0135)	(0.0134)
domestic							-0.000557	-0.000484	-0.000475	-0.000343
							(0.000518)	(0.000514)	(0.000531)	(0.000528)
hmt							0.000292	0.000651	0.000311	0.000712
							(0.000613)	(0.000626)	(0.000615)	(0.000628)
foreign							7.56e-05	0.000105	-4.06e-06	-2.33e-05
							(0.000818)	(0.000810)	(0.000826)	(0.000817)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-2.082***	-9.323***	-2.255**	-9.823***	-2.299***	-10.44***	-2.344***	-10.50***	-2.392***	-11.27***
	(0.781)	(3.255)	(0.872)	(3.327)	(0.874)	(3.355)	(0.797)	(3.489)	(0.801)	(3.551)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.853	0.856	0.857	0.860	0.858	0.861	0.857	0.861	0.858	0.861
Hausman test	0.0124	0.0122	0.5809	0.1841	-	0.0676	0.3341	0.0914	0.0886	0.6218

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

	Random Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2
logpop	0.0901*	1.270*	0.102**	1.186*	0.118**	1.392*	0.103*	1.330*	0.113**	1.531**
	(0.0509)	(0.700)	(0.0514)	(0.707)	(0.0537)	(0.718)	(0.0535)	(0.749)	(0.0549)	(0.758)
logpop2		-0.0730*		-0.0670		-0.0785*		-0.0754*		-0.0870*
		(0.0431)		(0.0434)		(0.0440)		(0.0457)		(0.0462)
loggdppercap			0.0150	0.0137	0.0834	0.0955*	0.0129	0.0130	0.0849	0.0978*
			(0.0114)	(0.0114)	(0.0535)	(0.0536)	(0.0118)	(0.0117)	(0.0536)	(0.0536)
loggdppercap2					-0.00928	-0.0111			-0.00979	-0.0115
					(0.00711)	(0.00714)			(0.00712)	(0.00713)
secondarygdp			0.000251	0.000246	0.000311	0.000317				
			(0.000487)	(0.000484)	(0.000487)	(0.000484)				
servicegdp			5.56e-05	3.03e-05	0.000168	0.000162				
			(0.000456)	(0.000453)	(0.000462)	(0.000459)				
logmaxtemp			-0.0342	-0.0361	-0.0350	-0.0373*	-0.0375*	-0.0384*	-0.0381*	-0.0393*
			(0.0224)	(0.0223)	(0.0224)	(0.0223)	(0.0225)	(0.0224)	(0.0224)	(0.0223)
logmintemp			0.000978	0.000151	0.00153	0.000639	0.00101	0.000407	0.000855	9.72e-05
			(0.0121)	(0.0120)	(0.0120)	(0.0120)	(0.0120)	(0.0120)	(0.0120)	(0.0119)
logavgpre			-0.0157	-0.0136	-0.0168	-0.0146	-0.0160	-0.0138	-0.0171	-0.0148
			(0.0110)	(0.0110)	(0.0110)	(0.0110)	(0.0110)	(0.0110)	(0.0110)	(0.0110)
logavghumi			0.0322	0.0284	0.0309	0.0261	0.0333	0.0269	0.0319	0.0241
			(0.0286)	(0.0287)	(0.0286)	(0.0286)	(0.0290)	(0.0292)	(0.0290)	(0.0292)
logavgsun			0.0229*	0.0236*	0.0247*	0.0259**	0.0247*	0.0261**	0.0263**	0.0283**
			(0.0127)	(0.0127)	(0.0128)	(0.0127)	(0.0127)	(0.0127)	(0.0127)	(0.0127)
domestic							-0.000460	-0.000400	-0.000425	-0.000350
							(0.000511)	(0.000508)	(0.000509)	(0.000506)
hmt							0.000471	0.000686	0.000447	0.000692
							(0.000611)	(0.000622)	(0.000608)	(0.000618)
foreign							-8.30e-05	-1.51e-05	-0.000340	-0.000312
							(0.000803)	(0.000798)	(0.000822)	(0.000815)
Constant	-3.968***	-8.695***	-4.180***	-8.510***	-4.435***	-9.570***	-4.130***	-9.076***	-4.314***	-10.06***
	(0.415)	(2.838)	(0.444)	(2.863)	(0.497)	(2.937)	(0.446)	(3.056)	(0.477)	(3.114)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

Table 6.6_B City Size- Urban Area (Built-Up Area)

Fixed Effects										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2
logbuiltup	0.0470***	-0.0163	0.0479***	-0.00610	0.0508***	-0.00935	0.0481***	-0.0206	0.0503***	-0.0250
	(0.0128)	(0.0763)	(0.0129)	(0.0778)	(0.0130)	(0.0776)	(0.0130)	(0.0792)	(0.0131)	(0.0792)
logbuiltup2		0.00955		0.00816		0.00909		0.0104		0.0114
		(0.0114)		(0.0116)		(0.0116)		(0.0119)		(0.0119)
loggdppercap			0.0155	0.0166	0.109*	0.112*	0.0156	0.0165	0.0927	0.0977
			(0.0123)	(0.0124)	(0.0623)	(0.0625)	(0.0125)	(0.0125)	(0.0644)	(0.0647)
loggdppercap2					-0.0134	-0.0137			-0.0111	-0.0116
					(0.00877)	(0.00878)			(0.00908)	(0.00910)
secondarygdp			0.000423	0.000394	0.000468	0.000438				
			(0.000490)	(0.000492)	(0.000489)	(0.000491)				
servicegdp			0.000304	0.000301	0.000314	0.000311				
			(0.000464)	(0.000464)	(0.000462)	(0.000463)				
logmaxtemp			-0.0397*	-0.0391*	-0.0405*	-0.0398*	-0.0432*	-0.0430*	-0.0432*	-0.0430*
			(0.0232)	(0.0233)	(0.0232)	(0.0232)	(0.0235)	(0.0235)	(0.0235)	(0.0235)
logmintemp			0.00760	0.00777	0.00684	0.00701	0.00682	0.00685	0.00581	0.00578
			(0.0121)	(0.0121)	(0.0121)	(0.0121)	(0.0120)	(0.0120)	(0.0120)	(0.0120)
logavgpre			-0.0160	-0.0157	-0.0173	-0.0171	-0.0161	-0.0158	-0.0173	-0.0172
			(0.0109)	(0.0109)	(0.0109)	(0.0109)	(0.0109)	(0.0109)	(0.0110)	(0.0110)
logavghumi			0.0305	0.0306	0.0305	0.0306	0.0341	0.0351	0.0327	0.0338
			(0.0310)	(0.0311)	(0.0310)	(0.0310)	(0.0314)	(0.0315)	(0.0314)	(0.0314)
logavgsun			0.0225*	0.0223*	0.0247*	0.0244*	0.0226*	0.0219*	0.0248*	0.0241*
			(0.0129)	(0.0130)	(0.0130)	(0.0130)	(0.0130)	(0.0130)	(0.0131)	(0.0131)
domestic							-0.000576	-0.000577	-0.000437	-0.000430
							(0.000504)	(0.000504)	(0.000517)	(0.000517)
hmt							6.00e-05	-3.21e-05	8.22e-05	-1.78e-05
							(0.000601)	(0.000610)	(0.000600)	(0.000609)
foreign							5.22e-05	-3.07e-05	-8.53e-05	-0.000184
							(0.000797)	(0.000803)	(0.000804)	(0.000810)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-2.924***	-2.819***	-3.052***	-2.968***	-3.199***	-3.109***	-2.943***	-2.831***	-3.060***	-2.943***
	(0.0613)	(0.139)	(0.177)	(0.213)	(0.201)	(0.231)	(0.117)	(0.173)	(0.151)	(0.193)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.860	0.860	0.864	0.864	0.865	0.866	0.865	0.865	0.865	0.866
Hausman test	0.7844	0.4583	0.4890	0.6105	0.6651	0.7336	0.9999	0.0000	0.0064	-

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

Random Effects										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2
logbuiltup	0.0459***	0.0210	0.0450***	0.0349	0.0500***	0.00631	0.0442***	0.0410	0.0497***	0.00177
	(0.0122)	(0.0699)	(0.0124)	(0.0712)	(0.0127)	(0.0727)	(0.0126)	(0.0720)	(0.0128)	(0.0740)
logbuiltup2		0.00369		0.00149		0.00657		0.000477		0.00722
		(0.0102)		(0.0104)		(0.0108)		(0.0106)		(0.0110)
loggdppercap			0.0147	0.0146	0.102**	0.110**	0.0108	0.0108	0.108**	0.118**
			(0.0108)	(0.0108)	(0.0515)	(0.0530)	(0.0115)	(0.0115)	(0.0524)	(0.0545)
loggdppercap2					-0.0118*	-0.0129*			-0.0133*	-0.0147**
					(0.00679)	(0.00702)			(0.00697)	(0.00730)
secondarygdp			0.000291	0.000286	0.000365	0.000349				
			(0.000471)	(0.000472)	(0.000471)	(0.000472)				
servicegdp			4.95e-05	4.39e-05	0.000186	0.000174				
			(0.000438)	(0.000440)	(0.000444)	(0.000444)				
logmaxtemp			-0.0336	-0.0340	-0.0331	-0.0332	-0.0371*	-0.0373*	-0.0376*	-0.0377*
			(0.0217)	(0.0217)	(0.0216)	(0.0216)	(0.0220)	(0.0220)	(0.0219)	(0.0219)
logmintemp			0.000804	0.000788	0.00153	0.00144	0.000331	0.000347	3.48e-05	-0.000169
			(0.0116)	(0.0116)	(0.0116)	(0.0116)	(0.0117)	(0.0117)	(0.0116)	(0.0116)
logavgpre			-0.0161	-0.0160	-0.0175*	-0.0173	-0.0164	-0.0163	-0.0180*	-0.0179*
			(0.0106)	(0.0106)	(0.0106)	(0.0106)	(0.0107)	(0.0107)	(0.0107)	(0.0107)
logavghumi			0.0335	0.0340	0.0304	0.0306	0.0351	0.0354	0.0330	0.0333
			(0.0279)	(0.0279)	(0.0279)	(0.0279)	(0.0285)	(0.0285)	(0.0284)	(0.0284)
logavgsun			0.0198	0.0199	0.0216*	0.0217*	0.0211*	0.0211*	0.0229*	0.0230*
			(0.0123)	(0.0123)	(0.0123)	(0.0123)	(0.0124)	(0.0124)	(0.0123)	(0.0123)
domestic							-0.000477	-0.000473	-0.000433	-0.000394
							(0.000496)	(0.000498)	(0.000492)	(0.000495)
hmt							0.000193	0.000189	0.000129	7.74e-05
							(0.000594)	(0.000598)	(0.000590)	(0.000594)
foreign							-0.000115	-0.000106	-0.000477	-0.000516
							(0.000774)	(0.000775)	(0.000794)	(0.000797)
Constant	-3.439***	-3.404***	-3.542***	-3.526***	-3.731***	-3.680***	-3.459***	-3.455***	-3.620***	-3.568***
	(0.0809)	(0.127)	(0.175)	(0.207)	(0.206)	(0.223)	(0.122)	(0.159)	(0.148)	(0.168)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

Table 6.6_C City Size- Total GDP

	Fixed Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2
loggdp	-0.163**	-0.0135	-0.291***	-0.133	-0.286***	-0.133	-0.221**	0.0636	-0.216**	0.0674
	(0.0820)	(0.0869)	(0.0976)	(0.102)	(0.0979)	(0.102)	(0.0890)	(0.107)	(0.0901)	(0.108)
loggdp2		-0.0385***		-0.0386***		-0.0383***		-0.0517***		-0.0517***
		(0.00905)		(0.00925)		(0.00932)		(0.0116)		(0.0117)
loggdppercap			0.0204	0.00874	0.0677	0.0293	0.0190	0.00820	0.0394	0.0268
			(0.0125)	(0.0124)	(0.0626)	(0.0614)	(0.0127)	(0.0125)	(0.0656)	(0.0633)
loggdppercap2					-0.00679	-0.00294			-0.00293	-0.00267
					(0.00881)	(0.00859)			(0.00926)	(0.00893)
secondarygdp			0.000458	0.000495	0.000475	0.000502				
			(0.000496)	(0.000480)	(0.000497)	(0.000482)				
servicegdp			-0.000139	-0.000105	-0.000134	-0.000103				
			(0.000476)	(0.000461)	(0.000476)	(0.000462)				
logmaxtemp			-0.0339	-0.0427*	-0.0344	-0.0428*	-0.0405*	-0.0420*	-0.0405*	-0.0421*
			(0.0235)	(0.0229)	(0.0236)	(0.0229)	(0.0238)	(0.0230)	(0.0239)	(0.0230)
logmintemp			0.00716	0.00563	0.00682	0.00550	0.00850	0.00737	0.00825	0.00715
			(0.0122)	(0.0118)	(0.0122)	(0.0119)	(0.0121)	(0.0117)	(0.0122)	(0.0118)
logavgpre			-0.0121	-0.0103	-0.0127	-0.0106	-0.0132	-0.0108	-0.0136	-0.0111
			(0.0110)	(0.0107)	(0.0111)	(0.0107)	(0.0111)	(0.0107)	(0.0111)	(0.0108)
logavghumi			0.00681	0.01000	0.00724	0.0102	0.0159	0.0109	0.0158	0.0109
			(0.0324)	(0.0314)	(0.0324)	(0.0314)	(0.0326)	(0.0314)	(0.0326)	(0.0315)
logavgsun			0.0278**	0.0271**	0.0289**	0.0275**	0.0278**	0.0286**	0.0283**	0.0291**
			(0.0131)	(0.0127)	(0.0132)	(0.0128)	(0.0132)	(0.0127)	(0.0133)	(0.0129)
domestic							-0.000673	0.000187	-0.000634	0.000223
							(0.000514)	(0.000532)	(0.000529)	(0.000546)
hmt							2.55e-05	0.000958	4.01e-05	0.000971
							(0.000617)	(0.000631)	(0.000620)	(0.000633)
foreign							3.08e-05	0.000872	-4.60e-06	0.000840
							(0.000809)	(0.000802)	(0.000818)	(0.000811)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-2.196***	-2.112***	-1.779***	-1.652***	-1.869***	-1.692***	-2.028***	-2.215***	-2.074***	-2.256***
	(0.310)	(0.301)	(0.416)	(0.405)	(0.433)	(0.422)	(0.353)	(0.343)	(0.383)	(0.371)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.855	0.864	0.862	0.871	0.862	0.871	0.861	0.871	0.861	0.871
Hausman test	0.0245	0.0000	0.0023	0.0650	0.2240	-	-	-	0.0000	0.0000

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

	Random Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2
loggdp	0.0172	0.179***	0.0124	0.179***	0.0334	0.188***	0.00136	0.276***	0.0267	0.300***
	(0.0168)	(0.0504)	(0.0205)	(0.0559)	(0.0261)	(0.0566)	(0.0229)	(0.0777)	(0.0297)	(0.0798)
loggdp2		-0.0303***		-0.0300***		-0.0291***		-0.0433***		-0.0431***
		(0.00887)		(0.00930)		(0.00934)		(0.0117)		(0.0117)
loggdppercap			0.0168	0.00662	0.101	0.0660	0.0146	0.00455	0.1000	0.0878
			(0.0125)	(0.0127)	(0.0626)	(0.0623)	(0.0126)	(0.0127)	(0.0655)	(0.0642)
loggdppercap2					-0.0120	-0.00845			-0.0122	-0.0119
					(0.00881)	(0.00870)			(0.00920)	(0.00901)
secondarygdp			0.000192	0.000245	0.000229	0.000270				
			(0.000492)	(0.000482)	(0.000490)	(0.000481)				
servicegdp			-6.31e-05	4.41e-05	-8.92e-06	8.00e-05				
			(0.000465)	(0.000456)	(0.000464)	(0.000456)				
logmaxtemp			-0.0290	-0.0384*	-0.0305	-0.0394*	-0.0337	-0.0366*	-0.0336	-0.0363
			(0.0225)	(0.0222)	(0.0224)	(0.0222)	(0.0226)	(0.0221)	(0.0226)	(0.0221)
logmintemp			0.00171	0.000933	0.00171	0.000996	0.00189	0.00110	0.00136	0.000542
			(0.0122)	(0.0119)	(0.0121)	(0.0119)	(0.0122)	(0.0119)	(0.0122)	(0.0119)
logavgpre			-0.0157	-0.0143	-0.0168	-0.0151	-0.0161	-0.0143	-0.0174	-0.0156
			(0.0111)	(0.0109)	(0.0111)	(0.0109)	(0.0111)	(0.0109)	(0.0111)	(0.0109)
logavghumi			0.0296	0.0361	0.0303	0.0366	0.0314	0.0296	0.0303	0.0284
			(0.0289)	(0.0285)	(0.0288)	(0.0285)	(0.0292)	(0.0286)	(0.0292)	(0.0286)
logavgsun			0.0203	0.0197	0.0223*	0.0211*	0.0220*	0.0229*	0.0243*	0.0251**
			(0.0128)	(0.0125)	(0.0128)	(0.0126)	(0.0128)	(0.0125)	(0.0129)	(0.0126)
domestic							-0.000552	0.000227	-0.000387	0.000387
							(0.000528)	(0.000558)	(0.000542)	(0.000571)
hmt							0.000372	0.00114*	0.000441	0.00121*
							(0.000622)	(0.000643)	(0.000624)	(0.000645)
foreign							-0.000340	0.000388	-0.000477	0.000249
							(0.000829)	(0.000834)	(0.000835)	(0.000841)
Constant	-3.286***	-3.468***	-3.375***	-3.530***	-3.563***	-3.657***	-3.301***	-3.662***	-3.491***	-3.847***
	(0.0672)	(0.0863)	(0.173)	(0.177)	(0.222)	(0.220)	(0.127)	(0.159)	(0.191)	(0.211)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

Table 6.6_D City Size- Energy Consumption

	Fixed Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2
logelec	-0.00638	0.0284	-0.00818	0.0421	-0.00875	0.0411	-0.00837	0.0329	-0.00889	0.0324
	(0.00903)	(0.0420)	(0.00941)	(0.0429)	(0.00942)	(0.0428)	(0.00945)	(0.0432)	(0.00948)	(0.0432)
logelec2		-0.00498		-0.00721		-0.00714		-0.00595		-0.00594
		(0.00568)		(0.00582)		(0.00581)		(0.00585)		(0.00585)
logcoalgas	-0.00900	0.00883	-0.00609	0.00157	-0.00611	0.00943	-0.00649	0.00307	-0.00661	0.00855
	(0.00792)	(0.0305)	(0.00807)	(0.0313)	(0.00807)	(0.0318)	(0.00806)	(0.0313)	(0.00807)	(0.0317)
logcoalgas2		-0.00310		-0.00157		-0.00282		-0.00185		-0.00274
		(0.00461)		(0.00474)		(0.00482)		(0.00473)		(0.00482)
loglpg	-0.00692	0.0724*	-0.00987	0.0681	-0.0103	0.0706*	-0.00966	0.0807*	-0.00986	0.0811*
	(0.00980)	(0.0413)	(0.00990)	(0.0416)	(0.00991)	(0.0416)	(0.0101)	(0.0421)	(0.0101)	(0.0421)
loglpg2		-0.0117**		-0.0115**		-0.0120**		-0.0133**		-0.0134**
		(0.00572)		(0.00576)		(0.00577)		(0.00584)		(0.00585)
loggdppercap			0.0188	0.0209	0.0853	0.104	0.0193	0.0209	0.0708	0.0847
			(0.0128)	(0.0130)	(0.0638)	(0.0643)	(0.0130)	(0.0131)	(0.0660)	(0.0662)
loggdppercap2					-0.00951	-0.0120			-0.00737	-0.00919
					(0.00894)	(0.00908)			(0.00926)	(0.00935)
secondarygdp			0.000287	0.000277	0.000315	0.000313				
			(0.000503)	(0.000503)	(0.000504)	(0.000503)				
servicegdp			0.000140	0.000132	0.000141	0.000140				
			(0.000475)	(0.000472)	(0.000475)	(0.000472)				
logmaxtemp			-0.0410*	-0.0394	-0.0417*	-0.0396	-0.0432*	-0.0421*	-0.0432*	-0.0416*
			(0.0242)	(0.0243)	(0.0242)	(0.0242)	(0.0245)	(0.0245)	(0.0245)	(0.0245)
logmintemp			0.00971	0.00921	0.00927	0.00861	0.00994	0.01000	0.00939	0.00926
			(0.0125)	(0.0124)	(0.0125)	(0.0124)	(0.0123)	(0.0123)	(0.0124)	(0.0123)
logavgpre			-0.0130	-0.0143	-0.0137	-0.0153	-0.0130	-0.0143	-0.0136	-0.0151
			(0.0114)	(0.0113)	(0.0114)	(0.0113)	(0.0114)	(0.0113)	(0.0114)	(0.0113)
logavghumi			0.0323	0.0327	0.0324	0.0315	0.0340	0.0357	0.0331	0.0337
			(0.0319)	(0.0321)	(0.0319)	(0.0321)	(0.0323)	(0.0324)	(0.0323)	(0.0325)
logavgsun			0.0237*	0.0249*	0.0253*	0.0272**	0.0242*	0.0248*	0.0257*	0.0269**
			(0.0133)	(0.0133)	(0.0134)	(0.0134)	(0.0133)	(0.0133)	(0.0135)	(0.0134)
domestic							-0.000561	-0.000723	-0.000468	-0.000614
							(0.000517)	(0.000522)	(0.000531)	(0.000534)
hmt							0.000287	0.000196	0.000310	0.000222
							(0.000618)	(0.000617)	(0.000620)	(0.000618)
foreign							6.74e-05	0.000198	-2.27e-05	9.51e-05
							(0.000819)	(0.000821)	(0.000827)	(0.000828)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-2.690***	-2.853***	-2.804***	-2.983***	-2.898***	-3.113***	-2.751***	-2.936***	-2.823***	-3.031***
	(0.0905)	(0.113)	(0.188)	(0.200)	(0.207)	(0.223)	(0.133)	(0.149)	(0.161)	(0.178)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.854	0.857	0.858	0.862	0.859	0.863	0.859	0.863	0.859	0.863
Hausman test	0.9960	0.9996	0.8881	0.9170	0.9162	0.9848	0.0608	0.0001	0.0000	0.1814

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

	Random Effects									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2	logno_2
logelec	-0.00623	0.0249	-0.00602	0.0411	-0.00726	0.0412	-0.00569	0.0273	-0.00686	0.0297
	(0.00877)	(0.0408)	(0.00904)	(0.0422)	(0.00911)	(0.0420)	(0.00927)	(0.0430)	(0.00927)	(0.0435)
logelec2		-0.00451		-0.00678		-0.00701		-0.00480		-0.00525
		(0.00553)		(0.00572)		(0.00569)		(0.00582)		(0.00590)
logcoalgas	-0.00889	0.00781	-0.00709	-0.00172	-0.00641	0.00460	-0.00805	-0.000205	-0.00714	0.00492
	(0.00761)	(0.0297)	(0.00781)	(0.0308)	(0.00784)	(0.0310)	(0.00788)	(0.0314)	(0.00788)	(0.0319)
logcoalgas2		-0.00297		-0.00125		-0.00212		-0.00162		-0.00228
		(0.00450)		(0.00468)		(0.00470)		(0.00476)		(0.00484)
loglpg	-0.00645	0.0753*	-0.00855	0.0704*	-0.0101	0.0725*	-0.00685	0.0812*	-0.00919	0.0796*
	(0.00940)	(0.0402)	(0.00959)	(0.0411)	(0.00971)	(0.0409)	(0.00970)	(0.0425)	(0.00984)	(0.0429)
loglpg2		-0.0119**		-0.0115**		-0.0121**		-0.0128**		-0.0130**
		(0.00557)		(0.00567)		(0.00566)		(0.00588)		(0.00594)
loggdppercap			0.0203*	0.0211*	0.0745	0.0913*	0.0160	0.0161	0.0798	0.0915*
			(0.0109)	(0.0111)	(0.0527)	(0.0527)	(0.0118)	(0.0119)	(0.0541)	(0.0552)
loggdppercap2					-0.00725	-0.00946			-0.00864	-0.0103
					(0.00691)	(0.00693)			(0.00715)	(0.00732)
secondarygdp			0.000183	0.000183	0.000224	0.000240				
			(0.000480)	(0.000480)	(0.000482)	(0.000480)				
servicegdp			-4.66e-05	-6.00e-05	2.96e-05	4.47e-05				
			(0.000447)	(0.000446)	(0.000453)	(0.000451)				
logmaxtemp			-0.0324	-0.0309	-0.0326	-0.0314	-0.0343	-0.0323	-0.0356	-0.0309
			(0.0224)	(0.0225)	(0.0224)	(0.0224)	(0.0228)	(0.0229)	(0.0227)	(0.0230)
logmintemp			0.00334	0.00355	0.00394	0.00436	0.00319	0.00403	0.00317	0.00365
			(0.0119)	(0.0118)	(0.0119)	(0.0118)	(0.0121)	(0.0121)	(0.0120)	(0.0122)
logavgpre			-0.0150	-0.0161	-0.0156	-0.0170	-0.0154	-0.0165	-0.0162	-0.0178
			(0.0109)	(0.0109)	(0.0109)	(0.0108)	(0.0111)	(0.0110)	(0.0110)	(0.0112)
logavghumi			0.0317	0.0318	0.0300	0.0295	0.0320	0.0320	0.0312	0.0283
			(0.0286)	(0.0289)	(0.0286)	(0.0289)	(0.0292)	(0.0295)	(0.0292)	(0.0296)
logavgsun			0.0206	0.0218*	0.0218*	0.0237*	0.0218*	0.0226*	0.0232*	0.0240*
			(0.0125)	(0.0125)	(0.0126)	(0.0125)	(0.0127)	(0.0126)	(0.0127)	(0.0128)
domestic							-0.000516	-0.000660	-0.000485	-0.000648
							(0.000507)	(0.000516)	(0.000505)	(0.000522)
hmt							0.000338	0.000282	0.000302	0.000245
							(0.000610)	(0.000612)	(0.000608)	(0.000619)
foreign							-0.000255	-8.11e-05	-0.000485	-0.000454
							(0.000791)	(0.000797)	(0.000815)	(0.000832)
Constant	-3.141***	-3.303***	-3.264***	-3.440***	-3.358***	-3.575***	-3.219***	-3.392***	-3.299***	-3.495***
	(0.0891)	(0.114)	(0.183)	(0.197)	(0.203)	(0.220)	(0.129)	(0.148)	(0.145)	(0.167)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

Table 6.6_E City Size by Energy Consumption- Decomposed to Residential Energy Use

Fixed Effects

VARIABLES	(1) logno_2	(2) logno_2	(3) logno_2	(4) logno_2	(5) logno_2	(6) logno_2	(7) logno_2	(8) logno_2	(9) logno_2	(10) logno_2
logresielec	0.0175*	0.00799	0.0192**	0.0191*	0.0175	0.0160	0.0190*	0.0190*	0.0103	0.00897
	(0.00953)	(0.0412)	(0.00971)	(0.00972)	(0.0427)	(0.0428)	(0.00991)	(0.00992)	(0.0429)	(0.0430)
logresielec2		0.00121			0.000178	0.000351			0.00116	0.00133
		(0.00557)			(0.00581)	(0.00581)			(0.00582)	(0.00583)
logcoalgasresi	-0.00803	0.0216	-0.00750	-0.00746	0.0167	0.0210	-0.00769	-0.00774	0.0142	0.0175
	(0.00790)	(0.0309)	(0.00812)	(0.00812)	(0.0312)	(0.0315)	(0.00806)	(0.00807)	(0.0312)	(0.0315)
logcoalgasresi2		-0.00447			-0.00360	-0.00428			-0.00326	-0.00376
		(0.00466)			(0.00471)	(0.00475)			(0.00470)	(0.00475)
loglpgresi	0.00718	-0.0200	0.00538	0.00514	-0.0273	-0.0281	0.00572	0.00572	-0.0211	-0.0237
	(0.00980)	(0.0414)	(0.00989)	(0.00990)	(0.0418)	(0.0418)	(0.0102)	(0.0102)	(0.0428)	(0.0430)
loglpgresi2		0.00381			0.00465	0.00471			0.00384	0.00420
		(0.00589)			(0.00594)	(0.00594)			(0.00611)	(0.00613)
loggdppercap			0.0139	0.0700	0.0137	0.0773	0.0135	0.0573	0.0126	0.0658
			(0.0126)	(0.0632)	(0.0131)	(0.0639)	(0.0128)	(0.0653)	(0.0133)	(0.0662)
loggdppercap2				-0.00804		-0.00915		-0.00629		-0.00766
				(0.00888)		(0.00900)		(0.00919)		(0.00933)
secondarygdp			0.000201	0.000226	0.000218	0.000245				
			(0.000504)	(0.000505)	(0.000509)	(0.000509)				
servicegdp			-1.26e-05	-1.10e-05	-2.01e-05	-2.09e-05				
			(0.000480)	(0.000480)	(0.000483)	(0.000483)				
logmaxtemp			-0.0366	-0.0371	-0.0366	-0.0371	-0.0390	-0.0390	-0.0388	-0.0387
			(0.0238)	(0.0238)	(0.0239)	(0.0239)	(0.0241)	(0.0241)	(0.0242)	(0.0242)
logmintemp			0.00795	0.00752	0.00790	0.00743	0.00838	0.00785	0.00835	0.00773
			(0.0124)	(0.0124)	(0.0124)	(0.0124)	(0.0123)	(0.0123)	(0.0124)	(0.0124)
logavgpre			-0.0159	-0.0166	-0.0149	-0.0156	-0.0161	-0.0168	-0.0154	-0.0162
			(0.0111)	(0.0112)	(0.0112)	(0.0113)	(0.0112)	(0.0112)	(0.0112)	(0.0113)
logavghumi			0.0263	0.0264	0.0256	0.0256	0.0275	0.0267	0.0264	0.0253
			(0.0320)	(0.0320)	(0.0321)	(0.0321)	(0.0324)	(0.0324)	(0.0326)	(0.0326)
logavgsun			0.0250*	0.0263*	0.0245*	0.0260*	0.0258*	0.0270**	0.0256*	0.0271**
			(0.0133)	(0.0134)	(0.0134)	(0.0135)	(0.0134)	(0.0135)	(0.0134)	(0.0136)
domestic							-0.000537	-0.000457	-0.000491	-0.000391
							(0.000515)	(0.000528)	(0.000526)	(0.000540)
hmt							0.000376	0.000395	0.000404	0.000430
							(0.000620)	(0.000621)	(0.000629)	(0.000630)
foreign							-0.000285	-0.000362	-0.000352	-0.000452
							(0.000834)	(0.000842)	(0.000848)	(0.000857)
City Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-2.865***	-2.832***	-2.919***	-3.001***	-2.883***	-2.976***	-2.899***	-2.963***	-2.861***	-2.935***
	(0.0828)	(0.109)	(0.188)	(0.209)	(0.208)	(0.227)	(0.131)	(0.161)	(0.153)	(0.178)
Observations	300	300	299	299	299	299	299	299	299	299
R-squared	0.855	0.856	0.860	0.860	0.860	0.861	0.860	0.860	0.861	0.861
Hausman test	0.7439	0.8635	0.6325	-	0.0015	0.9247	0.0000	0.0005	0.0319	1.0000

"-" model fitted on these data fails to meet the asymptotic assumption of the Hauman test.

Random Effects

VARIABLES	(1) logno_2	(2) logno_2	(3) logno_2	(4) logno_2	(5) logno_2	(6) logno_2	(7) logno_2	(8) logno_2	(9) logno_2	(10) logno_2
logresielec	0.0155*	0.00289	0.0167*	0.0166*	0.0150	0.0165	0.0167*	0.0174*	0.00167	0.00474
	(0.00933)	(0.0398)	(0.00955)	(0.00959)	(0.0416)	(0.0419)	(0.00988)	(0.00986)	(0.0421)	(0.0419)
logresielec2		0.00160			0.000140	-0.000131			0.00201	0.00166
		(0.00537)			(0.00563)	(0.00567)			(0.00572)	(0.00568)
logcoalgasresi	-0.00879	0.0205	-0.00946	-0.00950	0.0131	0.0182	-0.00924	-0.00964	0.00822	0.0153
	(0.00772)	(0.0295)	(0.00797)	(0.00800)	(0.0304)	(0.0311)	(0.00805)	(0.00803)	(0.0309)	(0.0312)
logcoalgasresi2		-0.00448			-0.00346	-0.00427			-0.00264	-0.00378
		(0.00444)			(0.00461)	(0.00471)			(0.00466)	(0.00472)
loglpgresi	0.00811	-0.00718	0.00621	0.00636	-0.0133	-0.0161	0.00602	0.00559	-0.00453	-0.0120
	(0.00951)	(0.0402)	(0.00971)	(0.00975)	(0.0412)	(0.0416)	(0.0102)	(0.0102)	(0.0428)	(0.0429)
loglpgresi2		0.00204			0.00272	0.00314			0.00148	0.00246
		(0.00570)			(0.00583)	(0.00589)			(0.00610)	(0.00611)
loggdppercap			0.0179	0.0581	0.0169	0.0662	0.0117	0.0718	0.0103	0.0794
			(0.0110)	(0.0517)	(0.0113)	(0.0530)	(0.0119)	(0.0531)	(0.0122)	(0.0542)
secondarygdp			9.26e-05	0.000123	9.27e-05	0.000132				
			(0.000483)	(0.000486)	(0.000489)	(0.000494)				
servicegdp			-0.000177	-0.000118	-0.000200	-0.000131				
			(0.000454)	(0.000462)	(0.000458)	(0.000466)				
logmaxtemp			-0.0301	-0.0291	-0.0293	-0.0278	-0.0327	-0.0331	-0.0324	-0.0329
			(0.0221)	(0.0221)	(0.0222)	(0.0223)	(0.0225)	(0.0225)	(0.0226)	(0.0226)
logmintemp			0.00218	0.00258	0.00193	0.00236	0.00174	0.00168	0.00187	0.00170
			(0.0119)	(0.0119)	(0.0120)	(0.0120)	(0.0121)	(0.0121)	(0.0122)	(0.0121)
logavgpre			-0.0166	-0.0172	-0.0159	-0.0165	-0.0171	-0.0181*	-0.0169	-0.0178
			(0.0108)	(0.0109)	(0.0109)	(0.0110)	(0.0110)	(0.0110)	(0.0111)	(0.0110)
logavghumi			0.0299	0.0278	0.0296	0.0268	0.0303	0.0290	0.0305	0.0290
			(0.0285)	(0.0286)	(0.0286)	(0.0287)	(0.0291)	(0.0291)	(0.0292)	(0.0292)
logavgsun			0.0212*	0.0220*	0.0204	0.0211*	0.0230*	0.0242*	0.0226*	0.0238*
			(0.0126)	(0.0127)	(0.0127)	(0.0128)	(0.0127)	(0.0127)	(0.0128)	(0.0128)
loggdppercap2				-0.00538		-0.00660		-0.00817		-0.00942
				(0.00678)		(0.00696)		(0.00703)		(0.00720)
domestic							-0.000513	-0.000488	-0.000499	-0.000450
							(0.000509)	(0.000508)	(0.000522)	(0.000519)
hmt							0.000457	0.000426	0.000484	0.000449
							(0.000618)	(0.000616)	(0.000625)	(0.000621)
foreign							-0.000611	-0.000850	-0.000612	-0.000881
							(0.000809)	(0.000837)	(0.000822)	(0.000849)
Constant	-3.310***	-3.297***	-3.380***	-3.457***	-3.364***	-3.461***	-3.346***	-3.432***	-3.322***	-3.422***
	(0.0898)	(0.111)	(0.181)	(0.206)	(0.199)	(0.224)	(0.130)	(0.149)	(0.150)	(0.167)
Observations	300	300	299	299	299	299	299	299	299	299
Number of city	30	30	30	30	30	30	30	30	30	30
LM test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Test for whether using random effect or simple OLS: Breusch-Pagan Lagrange Multiplier (LM).

6.5 CONCLUSION

Using data for 30 major cities in China from 2003 to 2012, we explore the relationship between local air quality (PM_{10}, SO_2, and NO_2) and city characteristics (city size, city industrial composition, city technique composition, and city climate conditions). This chapter provides the first major examination of the effect of city size on the environment in China and we measure city size in terms of city population size, urban built-up area size, total GDP size, and energy consumption size. In addition, we also split the total energy consumption into residential energy consumption and industrial energy consumption to explore the impact of urban growth on local air quality.

Specifically, to conclude, we illustrate the mixed results for our 3 air quality pollutants PM_{10}, SO_2, NO_2 in Table 6.7. There seems to be an inverted-U shape between city population size and PM_{10} and NO_2 concentrations. City population size tends to be negatively related to SO_2 concentration levels. Urban built-up area has an inverted-U shape correlation with PM_{10} and SO_2 and a linear positive correlation with NO_2.

Total GDP size consistently shows an inverted-U shape relationship with the three air pollutants PM_{10}, SO_2, NO_2. Total energy consumption shows that coal gas is significantly correlated with PM_{10} (inverted-U) and SO_2 (negative). LPG is significantly correlated with SO_2 (positive) and NO_2 (inverted-U). From the residential energy usage perspective, coal gas mainly positively affects PM_{10} concentrations, and residential electricity usage significantly positively affects NO_2 concentrations. With respect to SO_2 concentrations, an increase in residential electricity and LPG usage will increase SO_2 concentrations, while increasing the level of coal gas usage decreases SO_2 concentrations.

For the city level industrial composition effects on air quality, we find that the percentage of secondary industry and also the service industry are positively associated with the PM_{10} and SO_2 concentrations.

With respect to the role of output from domestic/ HMT / foreign firms on city local air quality, output of domestic firms consistently positively affects the PM_{10} concentration. Output of HMT firms tends to negatively affect PM_{10}, but positively affects NO_2 concentrations. Increasing the share of output of foreign firms seems to increase PM_{10} concentrations but decreases NO_2 concentrations.

Finally, city precipitation shows a significant role in mitigating the PM_{10} and SO_2 concentration levels.

In addition, we also reconfirm the EKC hypothesis for some estimation specifications. When studying the PM_{10} concentration levels, in Table 6.4_D the random effects specification shows a significant positive estimated coefficient for GDP per capita, and a significant negative estimated coefficient for the squared term of GDP per capita. Similarly in the regression of city size by residential energy consumption for SO_2, random effects results for GDP per capita also confirm the EKC hypothesis, and show an inverted-U shape between GDP per capita and air pollution (see Table 6.5_E random effects).

Table 6.7 Mixed results table.

	City Size					Composition Effect		Technique Effect			City Climate Conditions		
	Population size	Urban area	Total GDP	Total energy consumption	Residential energy consumption	Secondary industry percentage	Service industry percentage	Output of domestic firms	Output of HMT firms	Output of foreign firms	Precipitation	Max temperature	Sunshine hours
PM10	Inverted-U shape	Inverted-U shape	Inverted-U shape	Coal gas shows inverted-U shape	Coal gas positive	positive	positive	Consistently positive	“_”	positive	negative	N/A	N/A
SO2	negative	Inverted-U shape	Inverted- U shape	Coal gas negative LPG positive	Electricity positive LPG positive Coal gas negative	positive	positive	N/A	N/A	negative	negative	positive	N/A
NO2	Inverted-U shape	positive	Inverted- U shape	LPG shows inverted-U shape	Electricity ‘positive	N/A	N/A	N/A	positive	N/A	N/A	negative	positive

APPENDIX

Figure A6.1: Shanghai 2012

Source: US Environmental Protection Agency

Figure A6.2: Beijing 2012

Source: US Environmental Protection Agency

CHAPTER 7
CONCLUSION

7.1 SUMMARY OF RESULTS

This thesis carefully investigates three topics within the broad area of urban growth and transformation in China. The first topic is the city size distribution which is described by Zipf's law and Gibrat's law, discussed in Chapters 3 and 4. The second topic is the city growth pattern which is tested by city sequential growth theory, explored in Chapter 5. The third topic is the environmental impact of urban growth studied in Chapter 6. We provide a full summary of results in this chapter.

Chapter 3 examines the evolution of Chinese city size distribution by testing for the well-known regularity - Zipf's law - from 1879 to 2009 (with number of cities varying by year). We estimate the Zipf's exponent using repeated cross-sectional analysis and results suggest that from the end of the Qing Dynasty in 1879 to the modern period of 2009, Chinese city size distribution becomes increasingly even (equal), i.e. the Zipf's exponent is increasing in absolute value. In other words, the size inequality becomes smaller over this century and the disparity between large and small cities is diminishing. Notably, Zipf's law emerged in the middle of this process in the end of the 1980s and early 1990s, Specifically, Chinese city size distribution evolves from less even than Zipf's prediction (before 'Economic Reform'[64], 1983) to very close to Zipf's prediction (end of 1980s and early 1990s) then grows to more even than Zipf's law prediction, then relatively stable with a Pareto exponent of around -1.15.

[64] "Economic Reform" was launched in 1979, and became to be effective in early 1980s. In this paper, we use 'Economic Reform' refers to the year that it became effective, i.e. early stage of 1980s, such as 1983, 1984, 1985 etc.

The results contribute to previous empirical studies on Zipf's law. Basically, the city size distribution we found in China is consistent with the literature which shows that Zipf's law will emerge under homogeneous urban growth, but city size needs time to converge to Zipf's law (Gabaix, 1999). However, no previous literature has covered the long run evolution of Chinese cities.

In addition, we also examine the city size distribution for different administrative attributes, different regions, and different historical experiences. Firstly, we divide the whole sample into prefecture-level cities and county-level cities and find that prefecture-level cities have a similar distribution to the whole sample, while county-level cities have long been distributed less evenly than Zipf's law predicts (the disparity between large and small cities is quite large), but the Pareto exponent shows a trend to Zipf's law. This might be because in the beginning county-level cities were not the priority of policy makers (although there are some exceptions[65]). Secondly, Zipf's law is found in Eastern region cities with a panel OLS regression Pareto exponent of -1.029 which cannot be rejected as being equal to -1 (at 1% significance). Thirdly, Zipf's law is found in the group of 82 older cities from 1936 to 2009.

Chapter 4 studied Gibrat's law which is a possible explanation for Zipf's law (Gabaix, 1999). We find that Gibrat's law approximately holds for the whole sample during the 1980s to 2000s, but not strictly. This can be considered the explanation for Zipf's law which appears to hold in the late 1980s and early 1990s and is consistent with Gabaix (1999)'s explanation for Zipf's law that if various cities grow at the same mean and variance (i.e. under 'Gibrat's-mode' of urban growth) then Zipf's law will emerge at the steady state. Similarly as in

[65] For instance, as a county-level city Puning City (in Guangdong Province) ranked 35th in 2009 by population, which is higher than most of the prefecture-level cities.

Chapter 3, we divide the whole sample into several subsamples and obtain mixed results: (1) Gibrat's law is more supported in prefecture-level cities and rejected in county-level cities, which is also consistent with previous Zipf's results that Zipf's law showed in prefecture-level cities' distribution and never showed in county-level cities. (2) Gibrat's law is found in Eastern region cities, which is also consistent with Zipf's results for Eastern regions which show that the Pareto exponent equals to -1.029, i.e. Zipf's law holds. (3) from the historical cities sample, we can see that the youngest city group d of 294 cities (1984-2009) is the closest to Gibrat's law, while other older historical cities are slowly moving towards Gibrat's mode of urban growth.

Chapter 5 further analyses the urban growth in a dynamic manner by testing for the sequential city growth theory following Cuberes (2009). We use the same Chinese city-level population data from 1879 to 2009 as in Chapters 3 and 4. We find similar results as Cuberes (2011) - the empirical evidence of sequential city growth theory- if we use the same econometric method. We investigate whether the rank of these fast growing cities rises as time goes by, implying that early on in the process of urbanization, fast-growers are concentrated in large-sized cities and then as time passes, the fast-growers can be found in middle-sized cities and then relatively small cities and so on. In other words, cities grow in sequential order, with the largest ones being the first to develop.

However, the driving force of the increasing average rank of the fast-growers may come from the increasing number of cities each year. To control for this factor we test for different subsamples attempting to test the validity of this sequential growth theory with slower growth rate of number of cities over time, or even without the number of cities growing over time by using a fixed number of cities. Results show that cities existing before the establishment of the People's Republic of China (PRC) show a sequential growth (1890-1953[66]), however, cities existing after the PRC establishment do not show sequential city growth or even a converse sequential growth (small cities initially grow the fastest) for cities existing after Economic Reform (1984[67]-2010).

In additon, to control for the effect of increasingly sample size in each year, we also rank city relatively from $1/N$ to N/N, N being the number of cities in each year. In other words, Cuberes's average rank of 'fast-growers' ($Rank25$) has been scaled and becomes the relative average rank of 'fast-growers' ($Rank25/N$). The results of relative average rank of 'fast-growers' do not seem to support a sequential city growth pattern.

The second part of Chapter 5 is the examination of the age dependent sequential city growth pattern (Sanchez-Vidal *et al.* 2014). Using parametric analysis we obtain the result that there are differences in city growth rates according to the age of the city. In general, when a city is born, it has the lowest growth rate; however, as decades pass, as it matures its growth rate accelerates, which is contrary to Sanchez-Vidal *et al.* 2014. This might be because historical cities in China have long been attractive to migrants and hence grow the fastest, as they may have better amenities and better job opportunities than new cities.

[66] People's Republic of China (PRC) established in 1949, we use 1953 in our data to represent the early stage of PRC establishment due to the data availability.
[67] Economic Reform was launched in 1979, we use 1984 in our data to represent the early stage of Economic Reform due to the data availability.

In chapter 6, we extend our analysis of city size and city growth from the previous chapters to examine the environmental impact of urban growth. We provide the first major examination of the effect of city size on the environment in China, using 30 major cities in China for 10 years from 2003 to 2012. We have mixed results for three local air pollutants- PM_{10}, $S0_2$, NO_2 concentrations.

We measure city size in terms of city population size, urban built-up area size, total GDP size, and energy consumption size. Firstly, for population size, we find that city population size seems to have an inverted-U shaped relationship with PM_{10} and NO_2 concentrations; and a negative linear relationship with SO_2 concentration levels. The possible explanation is that when the population size is small, the increasing population would lead to more PM_{10} concentration[68] because of more consumption due to more people in a city; while when the population size gets over a certain level, PM_{10} concentration tends to decrease (showing the inverted-U shape). This might because of the environment policy controlling for large polluted cities. The increasingly obvious rising level of PM_{10} concentration may attract local government's attention, and environment policy controls such as the 'lottery system for vehicle registration', the license plate limitation on using private vehicles[69] on road in Beijing might mitigate the concentration levels of PM_{10} (the downward sloping part of inverted-U shape).

With respect to the relationship between population size and NO_2 , one can easily explain that when population increase in a city the NO_2 concentration level will increase. However, our

[68] As mentioned before, city level PM_{10} concentration mainly comes from human activities, such as the burning of fossil fuels in vehicles, power plants and industrial process. In developing countries, the main method of supplying for heating and energy is the combustion of coal, which is a large source of the PM_{10} concentration.

[69] Private vehicles can be used on the road only according to the even or odd number of the last digit on their license plate in order to control the number of vehicles on the road.

results show that in these 30 major cities in China, as population getting over some certain level, the NO_2 concentration decrease with population increase. This might because of the increasing population within a city makes people living in smaller houses with less heating and commute less. As the main source of city level NO_2 emission is the using of vehicles, coal combustion power plants, thus the more population the less NO_2 concentration shown.

For the negative linear relationship between population size and SO_2 concentration, the possible explanation is that the main source of SO_2 emission is the various industrial process, especially the metallurgy. As population consistently increases in a city, the city's pollution intensity industries tend to be crowded out.

Secondly, the urban built-up area has an inverted-U correlation with PM_{10} and SO_2 and a linear positive correlation with NO_2. This can be explained that the increase of urban area naturally will decrease the concentration of air pollutants. Thirdly, Total GDP size consistently shows an inverted-U shape relationship with the three air pollutants PM_{10}, SO_2, NO_2, which is consistent with the EKC hypothesis. With respect to the total energy consumption size, we find that the consumption of coal gas is significantly correlated with PM_{10} (inverted-U) and SO_2 (negative). LPG is significantly correlated with SO_2 (positive) and NO_2 (inverted-U).

With respect to the role of output from domestic/ HMT / foreign firms on city local air quality, output of domestic firms consistently has a positive relationship with PM_{10} concentrations. Output of HMT firms tends to negatively affect PM_{10}, but positively affects NO_2 concentrations. Increasing the share of output of foreign firms seems to increase PM_{10}

concentrations but decreases NO_2 concentrations. This indicates that domestic firms might still concentrate in pollution intensive industries. This might indicates that domestic firms can be improved such as learning more advanced and clean technology in terms of affecting the local air quality.

To be noted that, with respect to Chapter 3 and Chapter 4, how the Chinese unique policies affect the validity of Zipf's law and Gibrat's law has briefly discussed. To conclude, in fact as all these three policies are the shocks for all cities, technically, all of them do not affect the validity of Zipf's law or Gibrat's law. Because if the shocks are applied to all cities, then cities still grow under identical growth process, where Gibrat's law may hold and then Zipf's law may also hold in the steady state. However, 'Economic Reform' may indeed promote more evenly distributed cities as it improve the level of economic development and the level of income for all cities. This may lead to the decrease of differences between city sizes, as cities are more identical for people to live for a better live (better income[70]).

Although these policies may not affect the validity of Zipf's law or Gibrat's law, we still discuss these policies because we are investigating city size distribution within China, while these policies are unique in China compared with studies focusing on other countries.

Examining the link between these policies to the validity of Zipf's law and Gibrat's law is interesting, but also challenging as there are still no exact literature investigating the precise relationship between policies and the validity of Zipf's law and Gibrat's law. In this thesis, we simply suggest how these policies may affect these empirical regularities.

[70] We do not consider the environment quality as the determinants of location choice of residence at this stage to address the current question more clearly.

We mainly focus on the empirical status of city size distribution of Chinese cities, i.e. whether the Chinese city size distribution reaches a steady state (Zipf's law), if not, how far is Chinese city size distribution below or above the steady state (Zipf's law). This is important as the city size distribution is attracting increasingly attention in regional studies, i.e., if population spreads too much and the size of single city will be too small, then it is hard to take the advantage of scale economy, and hard to construct close connections among cities which may waste the infrastructure construction; in contrast if population concentrates only in a few large cities, this will lead diseconomies of scale and congestion, environment degradation, housing shortage, employment difficulties, etc.

However, as we simply imply the possible mechanism of how these policies may influence city size distribution, we have to acknowledging that one does not have enough evidence to claim that these effects are indeed in place. More work on the precise relationship between the policy and the validity of the Zipf's law and Gibrat's law could be done in the future.

7.2 FUTURE RESEARCH

With the rapid growth of cities in China, Zheng and Kahn (2013) predict that there are 300 million future urbanites need to be accommodated in cities. This indicates that urban planners and policy makers should be aware of the upcoming economic, social and environmental issues. Economic research is helpful in terms of maximizing the economic benefit and minimizing the potential environment or social cost. Here we show some future economic research which can base on our research in this thesis.

The city size distribution and city growth pattern in China has been examined in the first three main chapters (Chapter 3, 4 and 5), the city growth issue still requires further investigation. As from 1890 to 2009 city size distribution in China approached to Zipf's law and then grew over Zipf's law, researches on the following period in the future are needed to investigate whether the city size distribution would return to Zipf's law, given that the distribution as Zipf's prediction is the steady state if cities grow under the mode of Gibrat's law states. This might provide valuable academic contribution to the debate between Zipf's law and Gibrat's law and the investigation of city growth model.

More accurate linkage between Chinese unique policies and city size distribution would be very interesting. In this thesis, we mainly focus on exploring the status of city size distribution in China; only simply describe some possible influences of these policies to city size distribution. It would be also interesting to investigate the policy implications of these two regularities (Zipf’s law and Gibrat’s law), as there seems no literature cover this field.

Regarding to the environment consequences of city growth in Chapter 6, it would be helpful if the city-level data of private vehicles is available for Chinese cities. Because that the exhaust from private vehicles using in cities are becoming one of the most influential factor in the degradation of city local air quality. With the private vehicle data we can more precisely investigate the determinants of the urban local air quality, examine the weight of each factor affecting the city local air quality, and in order to provide empirically urban policy advices.

REFERENCES

Alonso, W. (1964). "Location and land use. Toward a general theory of land rent." Location and land use. Toward a general theory of land rent.

Alonso, W. (1980). "Five bell shapes in development." Papers in Regional Science 45(1): 5-16.

Anderson, G. and Y. Ge (2005). "The size distribution of Chinese cities." Regional Science and Urban Economics 35(6): 756-776.

Andreoni, J. and A. Levinson (2001). "The simple analytics of the environmental Kuznets curve." Journal of public economics 80(2): 269-286.

Arrow, K., et al. (2004). "Are we consuming too much?" Journal of Economic Perspectives: 147-172.

Artigas, M. S., et al. (2011). "Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function." Nature genetics 43(11): 1082-1090.

Au, C. and V. Henderson. (2006). "How migration restrictions limit agglomeration and productivity in China." Journal of Development Economics 80: 350-388

Auerbach, F. (1913). Das gesetz der bevölkerungskonzentration.

Axtell, R. and R. Florida (2001). Emergent cities: a microeconomic explanation for Zipf's law, Society for Computational Economics.

Baskin, C. W. (1966). "Central places in southern Germany." Trans, of Christaller (1933). Englewood Cliffs, NJ: Prentice-Hall.

Benguigui, L. and E. Blumenfeld-Lieberthal (2007). "A dynamic model for city size distribution beyond Zipf's law." Physica A: Statistical Mechanics and its Applications 384(2): 613-627.

Benguigui, L. and E. Blumenfeld-Lieberthal (2011). "The end of a paradigm: is Zipf's law universal?" Journal of geographical systems 13(1): 87-100.

Berk, D. E. V., et al. (2001). "Composite quasar spectra from the sloan digital sky survey." The Astronomical Journal 122(2): 549.

Berry, B. J. and A. Okulicz-Kozaryn (2012). "The city size distribution debate: Resolution for US urban regions and megalopolitan areas." Cities 29: S17-S23.

Black, D. and V. Henderson (1999). "A theory of urban growth." Journal of political economy 107(2): 252-284.

Black, D. and V. Henderson (2003). "Urban evolution in the USA." Journal of Economic Geography 3(4): 343-372.

Blank, A. and S. Solomon (2000). "Power laws in cities population, financial markets and internet sites (scaling in systems with a variable number of components)." Physica A: Statistical Mechanics and its Applications 287(1): 279-288.

Bosker, M., et al. (2008). "A century of shocks: the evolution of the German city size distribution 1925–1999." Regional Science and Urban Economics 38(4): 330-347.

Brakman, S., et al. (2004). "The strategic bombing of German cities during World War II and its impact on city growth." Journal of Economic Geography 4(2): 201-218.

Brakman, S., et al. (2001). An introduction to geographical economics: Trade, location and growth, Cambridge university press.

Brakman, S., et al. (1999). "The return of Zipf: towards a further understanding of the rank‐size distribution." Journal of Regional Science 39(1): 183-213.

Brezis, E. S. and P. R. Krugman (1997). "Technology and the life cycle of cities." Journal of Economic growth 2(4): 369-383.

Bunten, D. and M. E. Kahn (2014). The Impact of Emerging Climate Risks on Urban Real Estate Price Dynamics, National Bureau of Economic Research.

Champernowne, D. G. (1953). "A model of income distribution." The Economic Journal: 318-351.

Chang, L. T. (2009). Factory girls: From village to city in a changing China, Random House LLC.

Cheshire, P. (1999). "Cities in competition: articulating the gains from integration." Urban Studies 36(5/6): 843.

Clark, J. S. and J. C. Stabler (1991). "Gibrat's Law and the growth of Canadian cities." Urban Studies 28(4): 635-639.

Cole, M. A. and R. J. Elliott (2003). "Determining the trade–environment composition effect: the role of capital, labor and environmental regulations." Journal of Environmental Economics and Management 46(3): 363-383.

Cole, M. A. and R. J. Elliott (2005). "FDI and the capital intensity of “dirty” sectors: a missing piece of the pollution haven puzzle." Review of Development Economics 9(4): 530-548.

Cole, M. A., et al. (2008). "The environmental performance of firms: The role of foreign ownership, training, and experience." Ecological Economics 65(3): 538-546.

Cole, M. A., et al. (2011). "Growth, foreign direct investment, and the environment: evidence from Chinese cities." Journal of Regional Science 51(1): 121-138.

Córdoba, J.-C. (2008). "On the distribution of city sizes." Journal of urban economics 63(1): 177-197.

Córdoba, J. C. (2008). "A GENERALIZED GIBRAT'S LAW*." International economic review 49(4): 1463-1468.

Corominas-Murtra, B. and R. V. Solé (2010). "Universality of Zipf’s law." Physical Review E 82(1): 011102.

Costa, D. L. and M. E. Kahn (2003). "Civic engagement and community heterogeneity: An economist's perspective." Perspectives on politics 1(01): 103-111.

Cuberes, D. (2009). "A model of sequential city growth." The BE Journal of Macroeconomics 9(1).

Cuberes, D. (2011). "Sequential city growth: Empirical evidence." Journal of urban economics 69(2): 229-239.

Darné, O. and C. Diebolt (2005). Non-stationarity tests in macroeconomic time series. New Trends in Macroeconomics, Springer: 173-194.

Dasgupta, S., et al. (2002). "Confronting the environmental Kuznets curve." Journal of economic perspectives: 147-168.

Davis, D. R. and D. E. Weinstein (2001). Bones, bombs and break points: the geography of economic activity, National Bureau of Economic Research.

Davis, J. C. and J. V. Henderson (2003). "Evidence on the political economy of the urbanization process." Journal of urban economics 53(1): 98-125.

Dimou, M. and A. Schaffar (2009). "Urban hierarchies and city growth in the Balkans." Urban Studies.

Dimou, M., et al. (2008). "La croissance urbaine chinoise reconsidérée." Région et développement 26: 109-131.

Dobkins, L. H. and Y. M. Ioannides (2000). "Dynamic evolution of the US city size distribution." The Economics of Cities: 217-260.

Duranton, G. (2002). "City size distributions as a consequence of the growth process."

Duranton, G. (2006). "Some foundations for Zipf's law: Product proliferation and local spillovers." Regional Science and Urban Economics 36(4): 542-563.

Duranton, G. (2007). "Urban evolutions: The fast, the slow, and the still." The American Economic Review: 197-221.

Duranton, G. and D. Puga (2004). "Micro-foundations of urban agglomeration economies." Handbook of regional and urban economics 4: 2063-2117.

Eaton, J. and Z. Eckstein (1997). "Cities and growth: Theory and evidence from France and Japan." Regional Science and Urban Economics 27(4): 443-474.

Ebenstein, A. (2012). "The consequences of industrialization: evidence from water pollution and digestive cancers in China." Review of Economics and Statistics 94(1): 186-201.

Ebenstein, A., et al. (2011). Chemical fertilizer and migration in China, National Bureau of Economic Research.

Ederington, J., et al. (2005). "Footloose and pollution-free." Review of Economics and Statistics 87(1): 92-99.

Ederington, J. and J. Minier (2003). "Is environmental policy a secondary trade barrier? An empirical analysis." Canadian Journal of Economics/Revue canadienne d'économique 36(1): 137-154.

Eeckhout, J. (2004). "Gibrat's law for (all) cities." American Economic Review: 1429-1451.

Eeckhout, J. (2009). "Gibrat's law for (all) cities: Reply." The American Economic Review 99(4): 1676-1683.

Ehrlich, I. and G. S. Becker (1972). "Market insurance, self-insurance, and self-protection." The Journal of Political Economy: 623-648.

El-Shakhs, S. (1972). "Development, primacy, and systems of cities." The Journal of Developing Areas: 11-36.

Eskeland, G. S. and A. E. Harrison (2003). "Moving to greener pastures? Multinationals and the pollution haven hypothesis." Journal of development economics 70(1): 1-23.

Fujita, M. (1982). "Simulation of adaptive modification of the vestibulo-ocular reflex with an adaptive filter model of the cerebellum." Biological cybernetics 45(3): 207-214.

Fujita, S. (1978). "Biology of early gastric carcinoma." Pathology-Research and Practice 163(4): 297-309.

Fujita, T. T., et al. (1978). Manual of downburst identification for project NIMROD, University.

Gabaix, X. (1999). "Zipf's Law and the Growth of Cities." American Economic Review: 129-132.

Gabaix, X. (1999). "Zipf's law for cities: an explanation." Quarterly journal of Economics: 739-767.

Gabaix, X. and R. Ibragimov (2011). "Rank− 1/2: a simple way to improve the OLS estimation of tail exponents." Journal of Business & Economic Statistics 29(1): 24-39.

Gabaix, X. and Y. M. Ioannides (2004). "The evolution of city size distributions." Handbook of regional and urban economics 4: 2341-2378.

Gan, L., et al. (2006). "Is the Zipf law spurious in explaining city-size distributions?" Economics Letters 92(2): 256-262.

Garmestani, A. S., et al. (2007). "Departures from Gibrat's law, discontinuities and city size distributions." Urban Studies 44(10): 1997-2007.

Gibrat, R. (1931). Les inégalités économiques, Recueil Sirey.

Giesen, K. and J. Südekum (2010). "Zipf's law for cities in the regions and the country." Journal of Economic Geography: lbq019.

Giesen, K. and J. Suedekum (2012). "The size distribution across all'cities': A unifying approach."

Glaeser, E. L. and M. E. Kahn (2010). "The greenness of cities: carbon dioxide emissions and urban development." Journal of urban economics 67(3): 404-418.

Glaeser, E. L., et al. (1995). "Economic growth in a cross-section of cities." Journal of monetary economics 36(1): 117-143.

Glaeser, E. L. and J. M. Shapiro (2003). "Urban growth in the 1990s: Is city living back?" Journal of regional science 43: 139-165.

González-Val, R. (2011). "Deviations from Zipf's Law for American Cities An Empirical Examination." Urban Studies 48(5): 1017-1035.

González-Val, R., et al. (2013). "Gibrat's law for cities, growth regressions and sample size." Economics Letters 118(2): 367-369.

Grossman, G. M. (1995). "2 Pollution and growth: what do we know?" The economics of sustainable development: 19.

Grossman, G. M. and A. B. Krueger (1994). Economic growth and the environment, National Bureau of Economic Research.

Guérin-Pace, F. (1995). "Rank-size distribution and the process of urban growth." Urban Studies 32(3): 551-562.

Hansen, N. (1990). "Impacts of small-and intermediate-sized cities on population distribution: issues and responses." Regional Development Dialogue 11(1): 60.

Harbaugh, W. T., et al. (2002). "Reexamining the empirical evidence for an environmental Kuznets curve." Review of Economics and Statistics 84(3): 541-551.

Harris, J. R. and M. P. Todaro (1970). "Migration, unemployment and development: a two-sector analysis." The American economic review: 126-142.

Henderson, J. V. (1986). "Efficiency of resource usage and city size." Journal of urban economics 19(1): 47-70.

Henderson, J. V. and Y. M. Ioannides (1981). "Aspects of growth in a system of cities." Journal of urban economics 10(1): 117-139.

Henderson, J. V. and A. J. Venables (2009). "The dynamics of city formation." Review of Economic Dynamics 12(2): 233-254.

Henderson, J. V. and H. G. Wang (2007). "Urbanization and city growth: The role of institutions." Regional Science and Urban Economics 37(3): 283-313.

Henderson, V. (2003). "The urbanization process and economic growth: The so-what question." Journal of Economic growth 8(1): 47-71.

Hilton, F. H. and A. Levinson (1998). "Factoring the environmental Kuznets curve: evidence from automotive lead emissions." Journal of Environmental Economics and Management 35(2): 126-141.

Ioannides, Y. M. (1994). "Product differentiation and economic growth in a system of cities." Regional Science and Urban Economics 24(4): 461-484.

Ioannides, Y. M. and H. G. Overman (2003). "Zipf's law for cities: an empirical examination." Regional Science and Urban Economics 33(2): 127-137.

Ioannides, Y. M. and S. Skouras (2009). "Gibrat's Law for (All) Cities: A Rejoinder." Available at SSRN 1481254.

Jia, T. and B. Jiang (2010). "Measuring urban sprawl based on massive street nodes and the novel concept of natural cities." arXiv preprint arXiv:1010.0541.

Jiang, B. and T. Jia (2011). "Zipf's law for all the natural cities in the United States: a geospatial perspective." International Journal of Geographical Information Science 25(8): 1269-1281.

Jiang, B. and X. Liu (2012). "Scaling of geographic space from the perspective of city and field blocks and using volunteered geographic information." International Journal of Geographical Information Science 26(2): 215-229.

Junius, K. (1999). "Primacy and economic development: Bell shaped or parallel growth of cities." Journal of Economic Development 24(1): 1-22.

Kahn, M. E. (1999). "The silver lining of rust belt manufacturing decline." Journal of urban economics 46(3): 360-376.

Kahn, M. E. (2006). Green cities: urban growth and the environment, Cambridge Univ Press.

Kahn, M. E. (2010). New evidence on trends in the cost of urban agglomeration. Agglomeration Economics, University of Chicago Press: 339-354.

Kahn, M. E. and J. Schwartz (2008). "Urban air pollution progress despite sprawl: the “greening” of the vehicle fleet." Journal of urban economics 63(3): 775-787.

Kahn, M. E. and R. Walsh (2014). Cities and the Environment, National Bureau of Economic Research.

Kahn, R. (2003). "Follow-up report on the diagnosis of diabetes mellitus: The expert committee on the diagnosis and classifications of diabetes mellitus*." Diabetes care 26(11): 3160.

Kamecke, U. (1990). "Testing the rank size rule hypothesis with an efficient estimator." Journal of urban economics 27(2): 222-231.

Kanemoto, Y. (1980). "Theories of urban externalities."

Kratz, M. and S. I. Resnick (1996). "The QQ-estimator and heavy tails." Stochastic Models 12(4): 699-724.

Krugman, P. (1992). A dynamic spatial model, National Bureau of Economic Research.

Krugman, P. (1996). "Confronting the mystery of urban hierarchy." Journal of the Japanese and International economies 10(4): 399-418.

Leamer, E. E. and J. Levinsohn (1994). International trade theory: the evidence, National Bureau of Economic Research.

Levy, M. (2009). "Gibrat's law for (all) cities: Comment." The American Economic Review 99(4): 1672-1675.

Li, S. and D. Sui (2013). "Pareto's law and sample size: a case study of China's urban system 1984–2008." GeoJournal 78(4): 615-626.

List, J. A. and C. A. Gallet (1999). "The environmental Kuznets curve: does one size fit all?" Ecological Economics 31(3): 409-423.

Lotka, A. J. (1925). "Elements of physical biology."

Lucas, R. E. (1988). "On the mechanics of economic development." Journal of monetary economics 22(1): 3-42.

Ma Laurence, J. and G. MaCui (1987). "Administrative changes and urban population in China." Annals of the Association of American Geographers 77(3): 373-395.

Madden, C. H. (1956). "On some indications of stability in the growth of cities in the United States." Economic development and cultural change: 236-252.

Malacarne, L. C., et al. (2001). "q-Exponential distribution in urban agglomeration." Physical Review E 65(1): 017106.

Malevergne, Y., et al. (2011). "Testing the Pareto against the lognormal distributions with the uniformly most powerful unbiased test applied to the distribution of cities." Physical Review E 83(3): 036111.

Michaels, G., et al. (2008). "Urbanization and structural transformation."

Miyao, T. (1981). Dynamic analysis of the urban economy, Elsevier.

Muth, R. F. (1969). "CITIES AND HOUSING; THE SPATIAL PATTERN OF URBAN RESIDENTIAL LAND USE."

Newman, M. E. (2005). "Power laws, Pareto distributions and Zipf's law." Contemporary physics 46(5): 323-351.

Nitsch, V. (2005). "Zipf zipped." Journal of urban economics 57(1): 86-100.

Ofek, H. and Y. Merrill (1997). "Labor immobility and the formation of gender wage gaps in local markets." Economic Inquiry 35(1): 28.

Palivos, T. and P. Wang (1996). "Spatial agglomeration and endogenous growth." Regional Science and Urban Economics 26(6): 645-669.

Panayotou, T. (2000). Economic growth and the environment, Center for International Development at Harvard University.

Pareto, V. (1964). Cours d'économie politique, Librairie Droz.

Parr, J. B. and K. Suzuki (1973). "Settlement populations and the lognormal distribution." Urban Studies 10(3): 335-352.

Paulsen, K. (2012). "Yet even more evidence on the spatial size of cities: Urban spatial expansion in the US, 1980–2000." Regional Science and Urban Economics 42(4): 561-568.

Pearson, C. E. (1980). "Rank-size distributions and the analysis of prehistoric settlement systems." Journal of Anthropological Research: 453-462.

Pearson, K. (1900). "X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling." The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 50(302): 157-175.

Peng, G. (2010). "Zipf's law for Chinese cities: Rolling sample regressions." Physica A: Statistical Mechanics and its Applications 389(18): 3804-3813.

Pfaff, A. S. (1999). "What drives deforestation in the Brazilian Amazon?: Evidence from satellite and socioeconomic data." Journal of Environmental Economics and Management 37(1): 26-43.

Quandt, R. E. (1964). "STATISTICAL DISCRIMINATION AMONG ALTERNATIVE HYPOTHESES AND SOME ECONOMIC REGULARITIES+." Journal of Regional Science 5(2): 1-23.

Resende, M. (2004). "Gibrat's law and the growth of cities in Brazil: A panel data investigation." Urban Studies 41(8): 1537-1549.

Roback, J. (1982). "Wages, rents, and the quality of life." The Journal of Political Economy: 1257-1278.

Rose, A. K. (2005). Cities and countries, National Bureau of Economic Research.

Rosen, K. T. and M. Resnick (1980). "The size distribution of cities: an examination of the Pareto law and primacy." Journal of urban economics 8(2): 165-186.

Rosen, S. (1979). "Wage-based indexes of urban quality of life." Current issues in urban economics 3.

Rossi-Hansberg, E. and M. L. Wright (2007). "Urban structure and growth." The Review of Economic Studies 74(2): 597-624.

Rozenfeld, H. D., et al. (2008). "Laws of population growth." Proceedings of the National Academy of Sciences 105(48): 18702-18707.

Rozenfeld, H. D., et al. (2009). The area and population of cities: New insights from a different perspective on cities, National Bureau of Economic Research.

Rozman, G. (1990). "East Asian urbanization in the nineteenth century: comparisons with Europe."

Saichev, A. I. and Y. Malevergne (2009). Theory of Zipf's law and beyond, Springer Science & Business Media.

Sánchez-Vidal, M., et al. (2014). "Sequential city growth in the US: Does age matter?" Regional Science and Urban Economics 44: 29-37.

Schaffar, A. and M. Dimou (2012). "Rank-size city dynamics in China and India, 1981–2004." Regional Studies 46(6): 707-721.

Selden, T. M. and D. Song (1994). "Environmental quality and development: is there a Kuznets curve for air pollution emissions?" Journal of Environmental Economics and management 27(2): 147-162.

Selden, T. M. and D. Song (1995). "Neoclassical growth, the J curve for abatement, and the inverted U curve for pollution." Journal of Environmental Economics and management 29(2): 162-168.

Semboloni, F. (2008). "Hierarchy, cities size distribution and Zipf's law." The European Physical Journal B 63(3): 295-301.

Sharma, S. (2003). "Persistence and stability in city growth." Journal of urban economics 53(2): 300-320.

Shostak, S. L., et al. (1991). "The dipole moment of water. I. Dipole moments and hyperfine properties of H2O and HDO in the ground and excited vibrational states." The Journal of chemical physics 94(9): 5875-5882.

Simon, H. A. (1955). "On a class of skew distribution functions." Biometrika: 425-440.

Singer, H. W. (1936). "The" courbe des populations." A parallel to Pareto's Law." The Economic Journal: 254-263.

Song, S. and K. H. Zhang (2002). "Urbanisation and city size distribution in China." Urban Studies 39(12): 2317-2327.

Soo, K. T. (2005). "Zipf's Law for cities: a cross-country investigation." Regional Science and Urban Economics 35(3): 239-263.

Soo, K. T. (2007). "Zipf's Law and urban growth in Malaysia." Urban Studies 44(1): 1-14.

Soo, K. T. (2014). "Zipf, Gibrat and geography: evidence from China, India and Brazil." Papers in Regional Science 93(1): 159-181.

Steindl, J. (1965). Random processes and the growth of firms: A study of the Pareto law, Griffin London.

Stern, D. I. and M. S. Common (2001). "Is there an environmental Kuznets curve for sulfur?" Journal of Environmental Economics and Management 41(2): 162-178.

Stokey, N. L. (1998). "Are there limits to growth?" International economic review: 1-31.

Urzúa, C. M. (2000). "A simple and efficient test for Zipf's law." Economics Letters 66(3): 257-260.

Urzúa, C. M. (2011). "Testing for Zipf's law: A common pitfall." Economics Letters 112(3): 254-255.

Wackernagel, M., et al. (2002). "Tracking the ecological overshoot of the human economy." Proceedings of the national Academy of Sciences 99(14): 9266-9271.

Wang, J., et al. (2008). "Advancing sustainable urban development in China." Proceedings of the ICE-Municipal Engineer 161(1): 3-10.

Wang, Z. and J. Zhu (2013). "Evolution of China's City-Size Distribution: Empirical Evidence from 1949 to 2008." Chinese economy 46(1): 38-54.

Williamson, J. G. (1965). "Regional inequality and the process of national development: a description of the patterns." Economic development and cultural change: 1-84.

Wormer, P. E. and A. van der Avoird (2000). "Intermolecular potentials, internal motions, and spectra of van der Waals and hydrogen-bonded complexes." Chemical reviews 100(11): 4109-4144.

Xing, Y. and C. D. Kolstad (2002). "Do lax environmental regulations attract foreign investment?" Environmental and Resource Economics 21(1): 1-22.

Xu, Z. and N. Zhu (2008). "Urban growth determinants in China." Chinese economy 41(1): 7-35.

Xu, Z. and N. Zhu (2009). "City size distribution in China: are large cities dominant?" Urban Studies 46(10): 2159-2185.

Ye, X. and Y. Xie (2012). "Re-examination of Zipf's law and urban dynamic in China: a regional approach." The Annals of Regional Science 49(1): 135-156.

Yichun, X. and A. K. Dutt (1990). "Regional investment effectiveness and development levels in China." GeoJournal 20(4): 393-407.

Yixing, Z. and L. J. Ma (2003). "China's urbanization levels: reconstructing a baseline from the fifth population census." The China Quarterly 173: 176-196.

Zheng, S., et al. (2011). China's rising demand for" green cities": Evidence from cross-city real estate price hedonics, National Bureau of Economic Research.

Zheng, S., et al. (2014). "Real estate valuation and cross-boundary air pollution externalities: evidence from Chinese cities." The Journal of Real Estate Finance and Economics 48(3): 398-414.

Zheng, S. and M. E. Kahn (2013). "Understanding China's urban pollution dynamics." Journal of Economic Literature 51(3): 731-772.

Zheng, S., et al. (2010). "Towards a system of open cities in China: Home prices, FDI flows and air quality in 35 major cities." Regional Science and Urban Economics 40(1): 1-10.

Zheng, S., et al. (2010). "The greenness of China: household carbon dioxide emissions and urban development." Journal of Economic Geography: lbq031.

Zipf, G. K. (1935). "The psycho-biology of language."

Zipf, G. K. (1949). "Human behavior and the principle of least effort."