The Problem of Growth-Inequality Nexus: An Analysis Based on the Case of China since Reform and Opening Up

Vom Fachbereich Rechts- und Wirtschaftswissenschaften

genehmigte Dissertation zur Erlangungung des Doktorgrades

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Vorgelegt von

Chengzhi Xu, MSc in International Business (Economics),

geboren in Anhui, P.R.China

Referenten:

Prof. Dr. Volker Caspari

Prof. Dr. Michael Neugart

Tag der Einreichung: 18.06.2013

Tag der Disputation: 14.11.2013

Ort der Einreichung: Darmstadt

D17, 2013

DEDICATED TO MY PARENTS

Acknowledgments

When I was little, I frequently heard the word “Communism” in China. Nobody knew what it is but everybody talked about it. I also heard the concept of “common prosperity” very early. During a quite long time when I was growing up in China, I had the chances to think about “common prosperity” even before I began to learn economics in Beijing Foreign Studies University. And my economic thinking has never really left it behind.

For me “common prosperity” is an ideal economy in which everyone enjoys similar prosperity. In order to achieve this, economic growth and equal distribution are both required and thus the relationship between growth and distribution is an important issue that economists must understand. This is the topic that I try to explore systematically in the dissertation based on my understanding on the so-called “economic miracle” history of China during the past 30 years. I hope the dissertation can contribute to the establishment of a more equal and prosperous global economy in the future.

I am deeply grateful for this dissertation’s supervisor Prof. Dr. Volker Caspari who encouraged and instructed me to work on this mission. Before meeting him, I haven’t realized that the relationship between growth and distribution is a different topic from the topics of understanding growth and distribution individually. Prof. Caspari is not a dogmatic economist and doesn’t limit me in any economic school’s ideology. This kind of “freedom” leads me to rethink the appropriateness of current economic analytical frames that I should apply or establish. Without Prof. Caspari’s supports and instructions, this dissertation is not possible and I cannot move toward on the road of becoming an economist. I also remain indebted to Prof. Dr. Michael Neugart for having kindly accepted to be my second referee and offering me his suggestions on revising this dissertation.

I have been benefited greatly from Dr. Kwok Tong Soo who was the supervisor of my Master thesis in Lancaster University, UK. His enthusiastic and instructive comments

on my work helped me make the dissertation updated and better. Dr. Ctirad Slavik of Frankfurt University, Dr. Masashi Hoshino of Waseda University and several other anonymous referees also contributed to the revisions of the dissertation. I am indebted to my fiancée Dr. Xiangchan Zhu who not only gave me much life support but also benefited me through many inspiring conversations and discussions on mathematics.

I hope to express many thanks to Mr. Yushi Mao who encouraged me to expand my work on the "transaction cost" in the dissertation and who kindly discussed the growth-inequality nexus with me in Beijing and confirmed the value of the general findings of the dissertation. I also owe my thanks to Prof. Justin Yifu Lin who gave me many encouragements and kindly sent me his book *Demystifying the Chinese Economy* when I started to write the dissertation. Prof. V. N. Balasubramanyam also read the dissertation and thanks for his encouragement.

Besides, I want to thank Prof. Dr. Ingo Barens, Prof. Dr. Jens Krüger and Prof. Dr. Ralf Elbert who kindly agreed to attend my dissertation defense. I am indebted to the helps of Thomas Weidinger, Le Wang, Jing Zhao, Yanqiang Ma, Ines Balta, Dr. Sabine Eschenhof-Kammer and Gülay Dolanbay. I also hope to thank Mr. Moritz Dornauf who kindly invited me as a Taichi instructor in his esteemed school in Darmstadt. This not only financially supported my study but also realized my intention to broadcast Chinese traditional culture to the western world.

Finally but most importantly, I am deeply grateful to my parents living far away in China. Their love and supports make me have the possibility and passion to pursue my dream. No word can express my gratitude to them. This dissertation is especially dedicated to them.

Chengzhi Xu (徐诚直)

18.11.2013, Darmstadt, Germany

Abstract

The dissertation contributes to both development economics and the new institutional economics. It mainly concerns long-run economic analysis. We argued and showed that pure neoclassical models without considering institutional factors cannot correctly analyze the real economic development, esp. for a transitional economy. Meanwhile, we showed that more neoclassical analytical skills can be incorporated into institutional analysis. The marginal analysis, for example, is successfully adopted in this dissertation dealing with the concepts like transaction cost and rent which often appear in the new institutional economics. The combination of both analytical frames can generate more explaining power on the real economy.

The dissertation discusses the relationship between economic growth and income/wealth distribution. Several views on the problem of growth-inequality nexus have been established: firstly, we confirmed that institutional arrangements are crucial for economic performance. A good institutional arrangement will benefit both economic growth and income/wealth distribution in the long run; secondly, we enhanced some of the former neoclassical viewpoints on wealth inequality evolution during growth, arguing that the crucial roles of rent distribution and more equal dissemination of education should be noted for achieving "common prosperity"; finally, we pointed out that inequality has a negative relationship with economic growth in the long run. To the opposite, equality takes a positive effect on sustainable growth, especially for a transitional economy experiencing the tertiarisation process.

This dissertation is a comprehensive study on the growth-inequality nexus based on the economic performance of P. R. China since reform and opening up in 1979. Thus the dissertation is also a contribution to the understanding of the so-called "China's Economic Miracle". The dissertation challenges the prevalent optimistic perspective on China's sustainable growth and provides counterarguments. The analysis also leads to suggestions for China's future reform which will help the economy overcome the "middle-income trap" problem.

The dissertation's content is summarized as follows: in Chapter 1 we summarized the basic characteristics of economic growth and income/wealth inequality during this period of China as a background introduction of the study. Then we made an empirical research in Chapter 2 on explaining China's income inequality since 1990s with institutional analysis, pointing out that the tertiarisation process plays a crucial role for both income/wealth distribution and sustainable economic growth. We argued that there are mainly three institutional arrangements in China blocking the process of tertiarisation: China's political and cultural institutions which make this country a very high rent-seeking economy, the double-track economic system and the dual-sector economy dividing into rural and urban sectors. In Chapter 3 we built three models to deepen our analysis. In the 1st model we extended the Ramsey-Cass-Koopmans growth model to explain how inequality evolves with two representative agents representing "inside institution" ("*tizhinei*") and "outside institution" ("*tizhiwai*") economies in China. The 2nd model shows that the double-track economic system should be transferred into a single-track one in order to maximize the output. We finally established a generalized model and proved that in a rent-seeking economy excluding labor wage, when the rent distribution is more equal, the wealth distribution will be also more equal in the long run. We turned to build a new theoretical frame based on a newly defined "transaction cost" in Chapter 4 in order to present another angle to look at both the short-run and long-run economic size changes. We put forward the concept of "unit transaction cost" and explored its relationship with economic growth. The analysis in this chapter confirms and deepens the arguments of the new institutional economics on economic growth. In Chapter 5, we extended our analysis on rent-seeking space and discussed the effect of inequality on economic growth from the perspective of rent. Combining the arguments of Chapter 2, the negative long-run relationship between income/wealth inequality and economic growth is established. Finally, we discussed the conclusions and implications of the dissertation in Chapter 6. We argued that one doctrine of Marxism that the economic base determines the superstructure is strongly doubtful in the short run if we treat 30 years' development as a short-run historical view.

Contents

Chapter 1

China's Growth-Inequality Performance since Reform and Opening up

1.1The economic dream of China

China is a country with more than 5000 years' civilization and the legendary story of Chinese history shows it began with a very ideal society under the production level at that time. The chief representative and one of major creators of Chinese civilization Kong Zi (Confucius) who lived from 551 BC to 479 BC called his dreaming society "Da Tong" which means a world of great harmony. This "Da Tong" society, in Confucius' view, has once nearly realized at a very ancient time in China. This society is not only equal but also enjoys satisfying wealth and happiness. He said:

大道之行也，天下为公，选贤与能，讲信修睦。故人不独亲其亲，不独子其子，使老有所终，壮有所用，幼有所长，矜、寡、孤、独、废疾者皆有所养，男有分，女有归。货恶其弃于地也，不必藏于己；力恶其不出于身也，不必为己。是故谋闭而不兴，盗窃乱贼而不作，故外户而不闭，是谓大同。(《礼记》)

(At the time of realizing "Da Tong", the world is publicly owned by all the people. The people who have the highest moral level and ability will be elected to be the leader. Every one of the society is trustable and people live together in harmony. People not only treat their own relatives as their relatives and not only treat their children as their children, which will lead the old to enjoy the rest of their lives, will have the youth to contribute according to the talent and will let the little to get the good education. The old without family and the disabled can enjoy supports. All the people have satisfied jobs and happy marriages. People neither like to waste the wealth nor to save them. People are worrying that they cannot fully contribute their

ability and their contributions are not for themselves. Thus the tricks will be checked and the bad things like stealing or robbery won't appear, even that you can safely leave the door open at night. Such a society is called the “Da Tong” society. *Record of Rites*)

As argued above, we know that the dream of “Da Tong” contains much more than just materialistic civilization. Such a society requires quite high moral level as well. But concerning the economic argument, I have to say that the above ancient Chinese ideal is quite similar to the idea of common prosperity with distribution based on need proposed by a great German scholar Karl Marx in late 19th century. It is thus not strange to me that communism has a deep cultural root in China since the country has been calling for the “Da Tong” society for thousands of years. The creator of People's Republic of China Zedong Mao even said in the late 1950s that the “Da Tong” society is the ideal society the communists hope to build. However, it is quite ironic to see the reality: today China's inequality situation is one of the worst countries in the world, under the leadership of the Communist Party of China (CPC). But we should not blame this to CPC blindly because before 1980s, China is a quite equal economy when its Gini coefficient was just about 0.16 in town and cities. It is obvious that with the rather equal society in early 1980s, China does not like the “Da Tong” society at all. In 1979, the year China began to reform and open up, the economy was very poor with about 250 million rural populations living in absolute poverty (*China Economic Times*, 2007.5.31). Also, strictly speaking, it was not so equal as the Gini coefficient shows. The equal situation only lies among the common population. The government officials who enjoyed a lot of invisible benefits and conveniences lived much better lives, although their wage income was not much higher than the most other people. Another opposite case is United States which is the strongest country in the world and its people are very rich. However, during 1980 to 2005, more than 80 percent of total increase in Americans' income went to the top 1 percent. The Nobel laureate in Economics Paul Krugman labeled the post-1979 epoch of U.S. the “Great Divergence”. Economic history clearly shows that it is very difficult to realize the

perfect goal: common prosperity.

There are several points I hope to point out here: the institutions after the primitive society make the perfect or absolute equality impossible. The human institutions have a solid core that is power. Somebody has more power than somebody else and some organizations have more power than some else. This power inequality is promised by varieties of contracts which are a reflection of institution. In the so-called capitalism society, the laborers and entrepreneurs create wealth together but obey different distribution rules. This is already analyzed by Adam Smith (1776). Whether capitalism is "the end of history" as what Francis Fukuyama argued in 1992 is not our concern here. At least, in the current predictable reality, we cannot define common prosperity as absolute equality. That is to say, the dream of common prosperity should not be a Utopia whose varieties of forms existing in past literatures. Utopia is also a good dream but not the real dream. The dream we are talking about is something that can be realized.

In a dynamic economic system, the only possibility of realizing common prosperity is to update the distribution shares during economic growth. Without economic growth, the redistribution will not be Pareto-improvement. The deprivation of wealthy group will destroy the economic efficiency and turn out to be the China's situation before reform and opening up. The growth-inequality nexus is thus a very important economics problem for realizing common prosperity. We start our analysis from China's economic background before reform and opening up.

1.2The economic background before 1979

The end of Mao's (Zedong Mao, first Chairman of P. R. China) time in 1976 offered China a new opportunity to change. In December 1978, the Third Plenary Session of the 11th Central Committee of the Communist Party of China (CPC) symbolized the beginning of the time "reform and opening up". Xiaoping Deng (1904-1997)，the recognized "core leader" of the 2nd generation of CPC is treated as the chief designer of this great development strategy.

From 1949 to 1979, China was a pure planned economy and rather closed. At Mao's time, frequent political movements were part of Chinese people's life. This phenomenon reached the peak during 1966 to 1976 when the Great Culture Revolution took place. Such political movements influenced the economic performance, e.g. all the colleges and universities were shut down during the Great Culture Revolution which seriously damaged the human capital accumulation and much of the working time was used to doing political movements. During these 30 years, ironically, Chinese people's average living standard made little improvement, although Mao promised in 1957 that China would catch up with the developed countries like UK within 15 years. This, however, doesn't mean that China's GDP growth performance in this period was very low. On the opposite, China's average GDP growth rate from 1952 to 1978 reached 4.39% with a negative TFP (Total Factor Productivity) growth rate of -1.37% according to the calculations of Maddison (2007). Zhu (2012) reported the TFP during this period was -1.07. This means that the good performance of growth during this period came from the accumulation of productive factors and suffered a high rate of inefficiency. Maddison's calculations also showed that the low TFP growth mainly came from the capital productivity. This phenomenon can be explained by the strategy of prioritizing heavy industry adopted by Chinese government during this period (see Table 1.1). The private investment in this period's China can be hardly found.

Table 1.1 Sectoral share of investment in capital construction, 1952–78
(%)

	Agriculture	Light industries	Heavy industries	Other
First Five-Year Plan (1953–57)	7.1	6.4	36.2	50.3
Second Five-Year Plan (1958–62)	11.3	6.4	54.0	28.3
(1963–65)	17.6	3.9	45.9	32.6
Third Five-Year Plan (1966–70)	10.7	4.4	51.1	33.8
Fourth Five-Year Plan (1971–75)	9.8	5.8	49.6	34.8

Note: The table was chosen from Justin Yifu Lin (2011) *Demystifying the Chinese Economy,* 1st Edition, Cambridge University Press. Original Source: Statistics on China's investment in fixed assets

(1950–78), compiled by the Department of Statistics on China's Investment in Fixed Assets, National Bureau of Statistics of China.

Imitating the development strategy of former Soviet Union, Chinese government believed that the catch-up strategy by prioritizing heavy industry was the correct way to realize socialism effectively. Obviously, for an agricultural economy like China in 1950s, it is not possible to support this capital-intensive strategy. Thus the economic system was twisted including suppress the wage rate in order to reduce the cost (Lin, 2011a). The planned economic system was endogenously promoted by the development strategy since it is not possible for government to subsidize the heavy industry for over 30 years without a firm control on the national resource (Lin, Cai and Li, 1994). This planned and distorted economy is proved to be lack of efficiency in every case of economies in global economic history.

In 1955, the State Council of China formally issued the documents to establish the system of the urban-rural dichotomy in household registration ("Hukou"). The household registration system has a long historical tradition in China for thousands of years. But besides the reason of historical tradition, the urban-rural dichotomy is also the result of prioritizing heavy industry strategy since it is not possible for the capital-intensive plants with planned economic system to create enough jobs for laborers from rural migrants (see Table 1.2). This institution lasts till now and there are still many obstacles to abolish it completely. But since 1979, the system has been loosened step by step. This dual economy has a huge impact on China's long-run development which I will explain further in later chapters.

Table 1.2 Sectoral employment structure, 1952–78

(%)

Year	Agriculture	Industry	Other
1952	83.5	6.0	10.5
1965	81.6	6.4	12.0
1978	73.3	12.5	14.2

Note: The table was chosen from Justin Yifu Lin (2011) *Demystifying the Chinese Economy,* 1st Edition,

Cambridge University Press. Original Source: National Bureau of Statistics of China 1992.

With the above background, the income/wealth distribution within this period appeared with two basic characteristics: firstly, the income inequality problem was successfully limited. It is reported that the Gini coefficient in 1952 was 0.22 (Risso and Carrera, 2010) and the Gini coefficient of the whole country was less than 0.3 in the late 1970s and 1980s (Feng, 2004), specifically it is reported that the Gini coefficient in 1978 was 0.29 for the rural area and 0.16 for the cities and towns (Xu and Li, 2011); secondly, the per capita income level was kept quite low. According to the calculation by Maddison (2007), the Chinese per capita income in 1952 was 538 (1990 GK Yuan) and in 1978 it became 978 (1990 GK Yuan). 978 Yuan was very low, compared to Hong Kong, for instance, that its per capita income was 9277 GK Yuan that year. Actually, China's per capita GDP measured in purchasing power parity in 1980 was only 30% of the average level of Sub-Saharan African economies. Further, in 1981 around 84% of Chinese people were living in poverty (Lin, 2012).

1.3 Growth and distribution performance since 1979

Usually, the past 33-year reform and opening up in China is divided into three parts: 1979-1992, 1992-2001 and 2001 till now. The first period is the early economic reform period which started from and mainly focused on rural areas. In 1992, with the well-known southern trip made Chinese leader Xiaoping Deng, China set up the goal of establishing the socialism marketing economic system in CPC's 14th National Congress which symbolized that the reform was deepened. In 2001, China entered the World Trade Organization (WTO) which means the opening up reached a new level.

This 33 years' economic performance in China presented a completely different picture from the past. Firstly, its average GDP growth rate has reached about 10% (from 1979 to 2009, China's average annual growth was 9.9%) with the lowest growth rate (less than 2%) happened in 1989 with the Tiananmen political turmoil. IMF (2010) reported that China's real GDP per head had increased nearly thirtyfold

over the past 30 years, measured at purchasing power parity. Now the World Bank lists China as an upper middle income country with per capita GDP reached 4428 USD in 2010. In 1978, the per capita GDP was 210 USD (Lin, 2011a). As early as in 2004, the World Bank reported that China's poverty ratio reduced to 2.8% of the total population.

Secondly, China's inequality situation has shown a continuous upward trend during the 30 years. It is commonly accepted that Gini coefficient already passed 0.5 in 2010. Jiangxi University of Finance and Economics even reported the Gini coefficient of Chinese households' income reached 0.61 in 2010. The National Bureau of Statistics, P. R. China also published the official Gini index for the past 10 years in 2013 (see Table 1.3). The data quality, however, has been under serious doubt and debate in China. Nevertheless, we can still see that the inequality degree in China is rather high.

Table 1.3

Year	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Gini	0.479	0.473	0.485	0.487	0.484	0.491	0.49	0.481	0.477	0.474

Source: National Bureau of Statistics, P.R. China

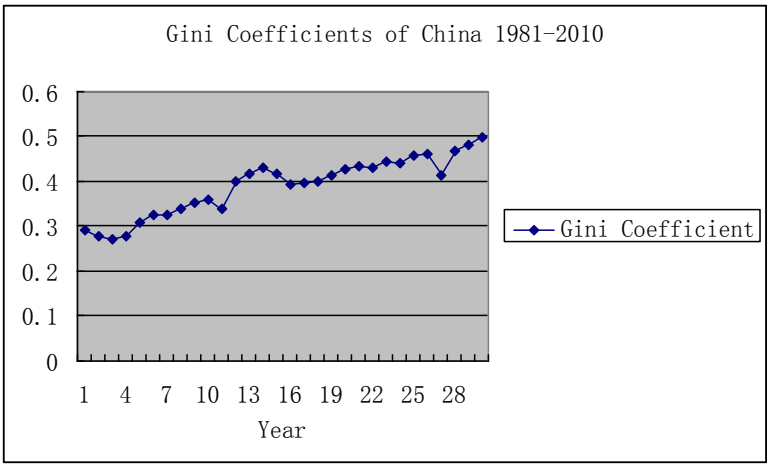
The comparative statistics between pre and post reform periods show that growth is more important than equality for overcoming poverty, at least for a poor economy like China in 1979. Also, the experience shows that higher inequality trend alone cannot stop the high growth rate or stagnate the GDP growth, at least for the early development period from a poor economy to a middle-income level economy.

Figure 1.1 and Figure 1.2 below describe the growth performance and income inequality performance in China from 1980 to 2010.

Figure 1.1

Note: Growth rates from 1980 to 1990 are calculated by GDP given by Maddison (2007) and growth rates from 1991 to 2010 are calculated by GDP given by International Monetary Fund, World Economic Outlook Database, April 2012.

Figure 1.2



Note: Gini coefficients from 1981-2004 (except 1991) are the results calculated by Cheng (2007); Gini coefficient for 1991 is the result calculated by Wang (2010); Gini indexes from 2005 to 2010 are chosen from several resources: Gini coefficients for 2005 and 2006 are chosen from Chen and Dai (2011). Gini indexes for 2007, 2008 and 2009 are reported by The World Factbook of Central Intelligence Agency, USA. It is a common sense that the Chinese Gini index in 2010 is above 0.5, according to a report made by Xinhua Agency published in May 21st, *China Economic Information* Daily. We set it as 0.5 here.

1.3.1 Basic Characteristics of China's growth performance since 1979

1. A gradual process of reform and opening up with a double-track economy

China's reform follows a way of gradually advanced reform which is opposite to the "shock therapy" adopted by Russia in early 1990s. This way means a double-track economy during the transition from a pure planned economy to a marketing economy under strong governmental control. Chinese government believes that this way can reduce the pressure against the reform. This is the basic characteristics of China's growth since 1979. The 30 years' experience proved that this process is at least successful for promoting growth. The opening up to the outside world adopts a gradual course. Private economy is allowed to develop but the traditional state-owned economy still plays a dominant role in the main strategic fields. The freedom of private economy also follows a gradual step. Government still keeps the right of deciding prices of many commodities, e.g. the oil and electricity price. And it also takes active measures with monetary and fiscal policies to intervene the economy, even in the prosperous normal time. As is shown in Figure 1.1, the growth performance in 1980s was much more volatile than that in the later periods. This phenomenon still lacks enough attention and explanation.

2. An urban-rural dual economy

As introduced earlier, the urban-rural dual economy was designed by Chinese government as early as in 1950s in order to support the development strategy of prioritizing heavy industry. This dual economic system keeps running during the reform and opening up. There is still 72% of Chinese population identified as rural "Hukou" holders (Meng, 2012). Although millions of rural laborers have been migrating to cities and towns, the registration institution ("Hukou") causes much discrimination to them in cities. They don't enjoy the equal treatment as urban citizens in many fields such as housing, insurance, children education, wage rate, etc. The reform on the urban-rural dual economy is rather complicated, facing high pressure from the urban side. Many of the Chinese economic problems are generated

by this specific institutional stickiness. A deeper analysis related to this will be presented in Chapter 2.

3. An economy with prevalent rent-seeking actions

China is an economy with huge rent-seeking space. According to Tollison (1982), rent-seeking activities are pointing to the activities of wasting resources competing for artificially contrived transfers. According to Hartle (1983), rent-seeking is to invest real resources in the expectation of gaining better situation of their wealth by changing legal rights or maximizing the benefit of earlier policy changes which created non-exclusive rights.

China is still far from a law-ruled society with high quality. Some people who are close to political power keep big advantages in the society in every aspect. The rights facing the law are different for different people/firms with "Guanxi" (meaning relationship) culture. With such a fact, China is an economy with prevalent rent-seeking activities. The so called "grey revenue" which is generated by the rents is huge in China and normal inequality calculation doesn't include this part. There are two levels of this kind of rent. The first comes from the double-track economic institutional arrangements in which the state-owned enterprises naturally keep a close relationship with governments. They are protected by the governments intentionally and thus easily access to financial market. It is reported that in 2010 more than 95% of the loans going to the SOEs. They also have many monopoly rights which mean that much of their revenue comes from the rent. The second level comes from the political institution which lacks efficient monitoring system. One of the leading principles of directing political reform is the "three insisting": insist on the leadership of Communist Party of China, insist on the democratic leadership of the people and insist on the rule of law. Obviously there are contradictions among the principles. In reality this means the government officials and their relatives or some "friends" can enjoy much better business opportunities and treatments. Such phenomenon is so prevalent in China that makes the whole economy a high-rent economy. Such social infrastructure makes a big influence during China's economic development which

obviously stays in the frame outside the neo-classical economics since there is no institutional factor in the neo-classical world. We will present a detailed analysis on this observation in the dissertation.

The above three points are the institutional characteristics of Chinese economy during growth. Concerning the pure growth performance, there are four other basic characteristics:

4. Mixed Pareto improvements

From the long-run perspective, both rural and urban Chinese people greatly benefited from the economic growth. But a more careful observation will find that the first period of the 30 years' reform (1980-1991) is a Pareto-improvement process and the second (1992-2000) and third period (2001-2010) are only to be a sense of "Kaldor-Hicks compensation" progress if we divide Chinese population into several groups. That means in the first period all groups of people in China were getting better but during the later two periods some groups of people were becoming worse, although this part of welfare loss can be compensated by the other parts. The Gini index cannot reflect this delicate difference.

In 1980s, with the beginning of Household Responsibility System (HRS)[1], the rural efficiency had greatly improved and thus the revenue increased. Meanwhile, as the beginning of city reform, the higher economic freedom brought the growth of the urban economy. During the 1980s, the reform of state-owned economy was staying at an early stage and workers were still enjoying stable and safe revenues. So there is no group of people suffered a loss during this period.

In 1990s, as the government promoted the reform of privatization of state-owned enterprises (SOE), many workers were laid off ("off-post"). The official jobless rate by the end of 1990s was over 10 million and the actual number was believed to be much higher (Gerth, 2010). These people gradually formed the poverty problem in cities and most of them still stay in poverty today.

5. Continuously growing investment

From the view of GDP accounting, in 2011, net export contributes zero to economic growth but the investment on capital-intensive sectors made by state-owned economies contributes more than 50% of the growth. And an obvious observation on China's economy is that the total investment amount in China from 1979 to 2009 has been continuously growing. The data of the real total investment in China is not available but we can use the "Total Fixed Assets Investment" made by the China National Statistics Bureau as a proxy of the trend. According to the statistics, the total fixed assets investment has been continuously rising during the past 30 years and the increasing speed became much quicker since 1990s. It is noted that the contribution ratio of the added value of the total fixed assets investment on GDP growth always surpasses the ratios of the other economic sectors' added values. Between 1978 and 2004 the average growth rate of the total fixed assets investment reached 12.4% (Wang and Cai, 2006). Also the state-owned investment has been always much larger than private sector's in China and the incremental speed of state-owned investment has been much faster, reflecting the characteristics of the socialism economy.

Figure 1.3

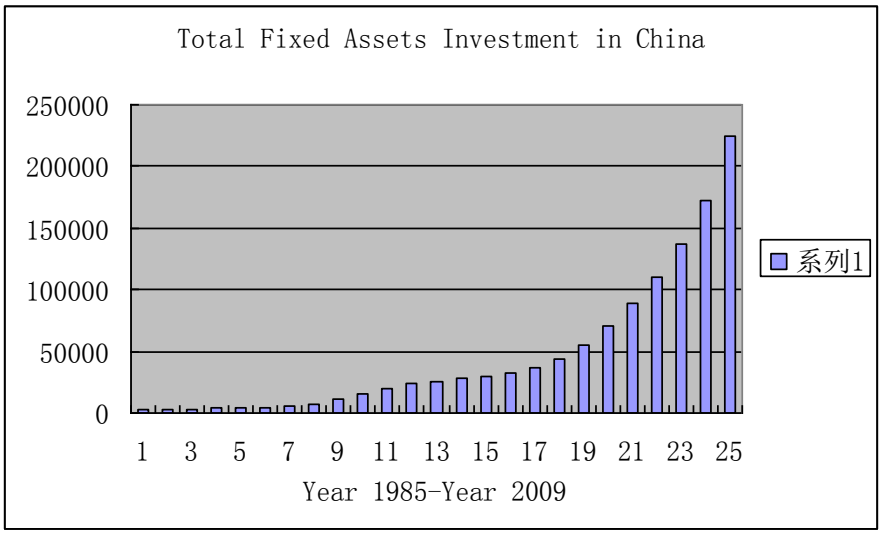
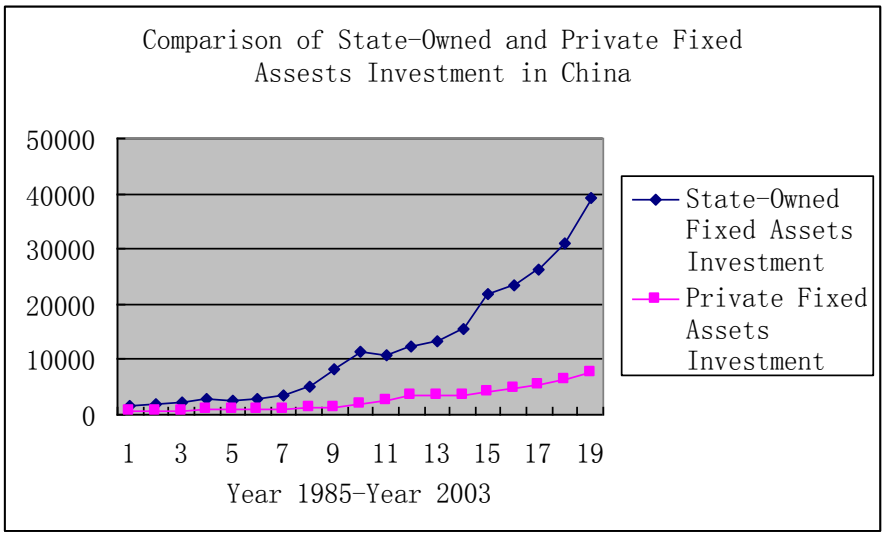


Figure 1.4



Unit: Hundred of Millions Yuan RMB

Resource: Data collected from the Annual Statistics Reports, National Bureau of Statistics, P. R. China

6. The composition of demand

The continuously growing supply ability in China has been supported by fast growing export with relatively low domestic consumption. According to the calculations of World Bank, China's export grew by 16% per year from 1979 to 2009. China's share in world exports of goods and non-factor services increased from 0.8% to 8.4% during the 30 years (Lin, 2011b). Meanwhile, China's international reserves reached 2130 billion USD in June 2009 (Song, Storesletten and Zilibotti (2011)).

China's household consumption as a percentage of GDP, however, fell from 51% in 1985 to 35% in 2009. The problem of insufficient domestic demand is caused by this which lowers down the speed of industrial update. We will also analyze this phenomenon in Chapter 2.

7. Industrial structure updates slowly

Although Chinese economic system since 1979 has gradually changed, the secondary industry still plays the biggest role in Chinese economic development. Global experience shows that it took around 20 years for East Asian "Four Dragon" economies to complete the industrial structural change from an economy mainly

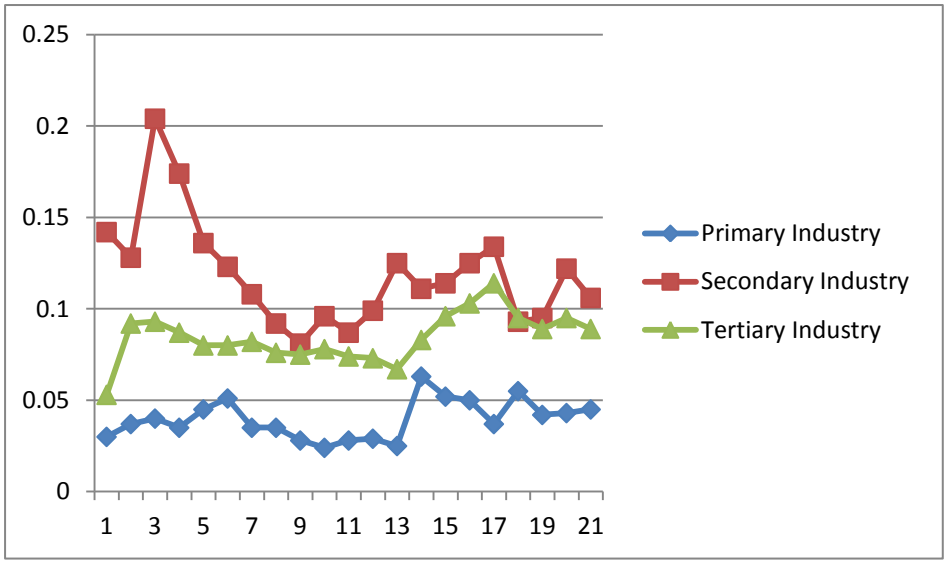
contributed by primary industry to the one contributed by tertiary (service) industry. China, however, is still struggling on climbing the "industry ladder" after 30 years' high growth experience.

Table 1.4 Industrial output growth rates between 1991 to 2011, China

Year	Primary Industry	Secondary Industry	Tertiary Industry
1991	0.03	0.142	0.053
1992	0.037	0.128	0.092
1993	0.04	0.204	0.093
1994	0.035	0.174	0.087
1995	0.045	0.136	0.08
1996	0.051	0.123	0.08
1997	0.035	0.108	0.082
1998	0.035	0.092	0.076
1999	0.028	0.081	0.075
2000	0.024	0.096	0.078
2001	0.028	0.087	0.074
2002	0.029	0.099	0.073
2003	0.025	0.125	0.067
2004	0.063	0.111	0.083
2005	0.052	0.114	0.096
2006	0.05	0.125	0.103
2007	0.037	0.134	0.114
2008	0.055	0.093	0.095
2009	0.042	0.095	0.089
2010	0.043	0.122	0.095
2011	0.045	0.106	0.089

Source: Data collected from the Annual Statistics Reports, National Bureau of Statistics, P. R. China

Figure 1.5 Industrial output growth rates between 1991 to 2011, China



Source: Data collected from the Annual Statistics Report s, National Bureau of Statistics, P. R. China

We can see from the above figure that during 1990s, there was a convergence of growth rates of secondary and tertiary industries. But this trend did not continue in 2000s. Generally the growth rates of secondary industry had always been higher than the other two industries. Till 2011, the percentage of added value of Chinese tertiary industry was still less than 50% which is even lower than the level of some other developing countries. But the added value of secondary industry reached 46.9% in 2011 which was above all historical achievements of developed countries.

1.3.2 Basic characteristics of China's income/wealth inequality since 1979

To be exact, income distribution is different from wealth distribution. But these two variables are often used as proxies to each other. We don't differentiate the two variables in the dissertation for the main goal is to see the growth-inequality nexus.

1. Inequality situation worsening in every aspect

China's inequality is mostly reflected in three aspects: rural-urban disparity, regional income disparity and income disparities among sectors. Data shows that for all these aspects the inequality situations have been worsening.

Table 1.5 China's Rural-Urban Income Disparity since 1979

Year	Per Capita net income of rural households (CNY)	Per capita disposable income of urban households (CNY)	Ratio of Incomes
1979	160	387	2.42
1980	191	477	2.50
1981	223	491	2.20
1982	270	526	1.95
1983	309	564	1.82
1984	355	651	1.83
1985	397	739	1.86
1986	423	899	2.12
1987	462	1002	2.17
1988	544	1181	2.17
1989	601	1375	2.29
1990	686	1510	2.20
1991	708	1700	2.40
1992	784	2026	2.58
1993	921	2577	2.80
1994	1221	3496	2.86
1995	1577	4283	2.71
1996	1926	4838	2.51
1997	2090	5160	2.47
1998	2162	5425	2.51
1999	2210	5854	2.65
2000	2253	6280	2.79
2001	2366	6859	2.90
2004	2936	9422	3.21
2005	3255	10493	3.22

2006	3587	11759	3.28
2007	4140	13786	3.33
2008	4761	15781	3.31
2009	5153	17175	3.33
2010	5919	19109	3.23

Source: *China Statistical Yearbook*, 2011, National Bureau of Statistics, China

Table 1.6 Comparison of eastern and western regions: rural per capita net income

Year	Eastern (CNY/person)	Western (CNY/person)	Ratio
1978	139	116	1.20
1990	848	534	1.59
2000	3063	1593	1.92

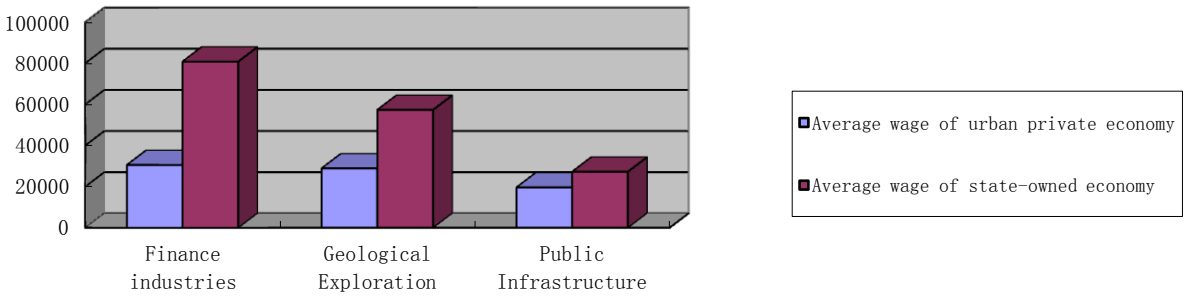
Source: Reports of the National Bureau of Statistics, China

Table 1.7 Comparison of eastern and western regions: urban per capita disposable income

Year	Eastern (CNY/person)	Western (CNY/person)	Ratio
1978	372	341	1.09
1990	1621	1265	1.28
2000	7682	5487	1.40

Source: Reports of the National Bureau of Statistics, China

Figure 1.6 Average wage difference between sectors in 2010



Source: Reports of the National Bureau of Statistics, China

It is reported that in 2010 the highest revenue industry in China is finance and the lowest is primary industry. The average wage difference between them reached 4.2: 1 and this ratio was between 1.6 -1.8:1 during 1980s (China People's Daily, 2011.12.2).

2. A strange but long-existing combination

If we divide China's economy into two industries of competitive private economy and monopolistic state-owned economy, we will find that the average wage is much higher in the state-owned industry (this can be seen in Figure 1.6). Further, its wage growth rate is also much higher. According to Banister (2007), the average real annual wage growth rate in unban manufacturing sector during 1992 to 2004 was 7.5% but if excluding state-owned and collectively owned economy, the growth rate becomes 4.6%. This is a unique phenomenon because the state-owned economy is lack of efficiency. There are large amount of studies showing that the productivity of Chinese private economy is much higher than that of the state-owned economy. A deeper observation shows that the private economy's income distribution is much more unequal. As a result, provinces with more private firms have much higher income inequality (Song, Storesletten and Zilibotti, 2011).

3. Who get what?

According to Gerth (2010), China's newly rich class can be divided into four groups: the individual entrepreneurs ("*getihu*"), people who took advantage of the

double-track price system, people who are land speculators and managers of state-owned enterprises. I would like to point out that there is another group obviously exits: the corrupted governmental officials of all levels. The poorest people are mainly the poorest in rural areas, most "off-post" workers of SOEs and most rural migrants in cities. It is obvious that the newly rich class, as introduced here, has a close relationship with rent-seeking activities under the economic and political institutions of China.

There is a comparatively rather small middle-income class in China. China now has a population close to 1.4 billion. Around 150 million own some luxury goods and 250 million can afford to buy some (Gerth, 2010). That means more than 1 billion people still cannot consider consuming some luxuries. Some reports (like Newspaper *Fazhiwanbao* July, 18th, 2010) also showed that the percentage of Chinese who became the middle-income class had reached 23%, if we put 11800-17700 U.S. dollars per year as the standard. The rather small middle-income class caused the inadequate inner effective demand for the productions. For the long time, Chinese economy is balanced by quick growth of exports. From 1979 to 2009, China's export growth rate reached 16% per year. When the coming of economic crisis happened in 2008, Chinese export sector also met with huge crisis since it is difficult for them to keep the business running with a shrunk global market. Chinese government recognized this problem and has been trying to expand the inner demand power with better distribution. So far this problem has not resolved and it becomes the major economic trouble for the sustainable growth of China.

4. Labor income share

Traditionally, income distribution is divided into two parts: the primary distribution and the redistribution. It is generally accepted that China's income inequality is mainly caused by the primary distribution (National School of Development, Peking University, 2013). Labor income share is an important part of the primary distribution. Functional distribution concerns the labor income share. Data shows that the labor income share had been keeping increasing since 1979 with the process of

decentralization of governmental economic powers. But this process ended by 1995 from when the labor income share has gradually reduced (China Development Report of People's Well-being 2011, pp.124-125). Such a reduction is prevalent in China except very few provinces or cities like Beijing.

Table 1.8 China GDP Structure (%)

Year	Labor Revenue	Net Taxation of Producers	Earned Surplus
1993	49.5	11.7	24.8
1994	50.3	12.0	23.8
1995	51.4	12.3	23.3
1996	51.2	12.9	23.2
1997	51.0	13.6	22.3
1998	50.8	14.3	21.7
1999	50.0	14.9	21.7
2000	48.7	15.3	21.9
2001	48.2	15.6	22.2
2002	47.8	15.8	22.9
2003	46.2	14.1	24.2
2004	41.6	14.1	30.3
2005	41.4	14.1	29.6
2006	40.6	14.6	30.7
2007	39.7	14.2	31.3

Source: Yearbooks of the National Bureau of Statistics, China

1.4 Concluding remarks

The above basic characteristics of China's growth-inequality performances since 1979 are the background for further theoretical and empirical studies. There are some other points we didn't include, e.g. production overcapacity during the past 10 years as a result of both insufficient domestic demand and blind investments. To fully explain all the details of these facts is not the task of this dissertation. We will focus on the relationship between China's growth experience and income/wealth distribution since 1979. But since all the economic factors are related, our analysis will also shed light on understanding these performances.

The Chinese economic development during the past 30 years is obviously unique and complicated but it also contains beneficial lessons for other developing economies, even for developed economies. The dissertation is a try to put forward a consistent logic contained in such historical economic facts with both empirical and theoretical studies.

Chapter 2

Inequality, Industrial Structure and Institutional Arrangements during China's Growth Miracle since 1990s

2.1 Introduction

This chapter studies the growth-inequality nexus based on China's economic performance since 1990s. There are several theoretical and empirical questions the chapter hopes to discuss. Firstly, China economy has already experienced more than 30 years' high growth rates since 1980s. Whether this performance can sustain and how long it can sustain generates many debates. To answer these important questions, we have to understand the crucial problems that block China's economic growth. This chapter tries to present a consistent analysis on the issue from the perspective of growth-inequality nexus. Secondly, the chapter aims to do an analysis on how the inequality is shaped during a dynamic structural change and how institutional arrangements are related.

The analysis will be based an empirical study on the influential factors of income inequality in China since 1990s. There are several considerations for choosing the time zone for analysis since 1990s: firstly, the 1980s for China is the early reform period which mainly focus on rural areas. Till 1990 when China completed "the 7th five-year plan", the tertiary industry (especially the privately-owned) had not really begun to develop. Thus if we include this period into the analysis, the effect of tertiary industry will be affected, which, as we will show, has a significant role in the problems we need to analyze. Secondly, we will have to face the problem of heterogeneous data quality since the only data available for this period is published by Chinese government which is different from our data set from 1990s.

This research is different from the static view on the determinants of income/wealth distribution since the process of economic growth is dynamic that nearly all the

variables of the economy will change during this long-run process. We are interested in how these changes affect the evolution of income/wealth distribution and which factor has the dominant effect. Further, the issue of income inequality in China should be paid more attention than usually thought for another reason: there are many so-called "style facts" of economic growth recognized in western economic performance. One of the facts is that in the long run balanced growth the functional income distribution (wL/rK) remains roughly constant (Kaldor, 1957). This is obvious unsettled in China since the Gini index shows that the income inequality has been continuously becoming larger (see Table 2.1 and Table 2.2). This may be or should be related to some other style facts which have not appeared in China either, like the Kuznets inverted U-shaped relation between income inequality and GDP per capita as well as the structural transformation from agriculture to service.

As expected, the empirical results of the chapter showed that the Kuznets "inverted U" curve has not appeared in China yet. Also, the chapter pointed out that the tertiarisation process (meaning the development of the tertiary industry comes to the stage of dominating the economy) is closely related to income distribution and the slow tertiarisation process has played the major role in explaining the increasing income inequality since 1990s in China. Then the chapter discussed the role of tertiarisation in the growth-distribution nexus theoretically and presented an institutional analysis on the phenomenon of slow tertiarisation in China afterwards. The chapter put forward that continuous reform on the "three" key institutional arrangements will be crucial for the sustainable development of China in the future: the political institution, the double-track economic institution and the urban-rural dual-sector institution. This research thus is meaningful for understanding the challenges faced by China to jump over the "middle-income trap".

The chapter is arranged as follows: section 2.2 is the literature review; section 2.3 presented the empirical study on explaining income inequality during China's growth experience since 1990s; section 2.4 discussed the role of tertiarisation in growth-inequality nexus; section 2.5 put forward an institutional analysis on the phenomenon of slow tertiarisation process of China; section 2.6 put forward several

policy suggestions for China' future reform; in section 2.7 we discussed the "pro-poor" growth problem based on China's experience; in section 2.8 a concluding remark is presented.

2.2 Literature review

The issue that how income inequality is influenced during economic growth is a very important economic problem and the economics field has been discussing it for quite a long time. It is a very interesting phenomenon that the classical economics tends to argue that the income inequality will become increasingly larger (see the review of Bronfenbrenner (1971) Chapter 4) whereas the neoclassical economics tends to support the opinion that the income inequality will become flattened with continuous economic growth (see Stiglitz (1969), Chatterjee (1994), Li, Xie and Zou (2000), etc.). The former stand is based on functional distribution theory and the later is usually applied with personal distribution theory. These two distribution theories, however, are not the curial cause for the ideas' difference. This debate hasn't reached a solid conclusion yet, although few economists continue to debate it. In the more recent economic history, the earliest well-known empirical and theoretical study on the problem of how income inequality is affected during economic growth comes from Kuznets (1955) whose discussion applies to both personal and functional distribution concepts.

Kuznets' research is based on the historical data for US, England and Germany. His argument leads to the well-known result that there is an "inverted U" relationship between national income and income inequality. This "inverted U" curve means that during the early period of development when the transition happens from the pre-industry to the industry civilizations, the income inequality widens; then the widening trend will become stabilized for some time and will reach the period of narrowing at the later stage of development. He also pointed out that there are at least two kinds of powers leading to greater inequality during the early development period. The first is the higher saving rate of richer people which brings about the difference of

assets' returning and the second is the structural change of an economy that development will bring the economy out of agriculture. With two basic facts that the average per capita income of rural population is lower than that of the urban and the urban areas often have larger inequality than the rural areas, he argued that with the increase of the urban population, the urban inequality will take higher effect. Also, the average productivity in urban areas increases much faster than that in rural areas. These are the reasons why early structural changes will lead to higher inequality.

For the power of saving's effect, Kuznets pointed out three counteracting factors: the demographic reason, the freedom of individual opportunity and the change of service income. Since the above three factors are influenced by different growth patterns, the final comprehensive result of the saving's power is indefinite. Here Kuznets emphasized the class mobility between the poor and the rich where the rich group doesn't have much rent so that the lower class' wealth can grow faster during the structural update.

Thus the "inverted U" curve mainly comes from the second power: the structural change. With a self-designed table of two economic sectors: agricultural and non-agricultural, Kuznets assumed reasonable numbers and calculated the results. The table he created contains much information on the evolution of income distribution and this table shows an "inverted U" curve where the inequality will be narrowed during the later stage of urbanization and industrialization. Kuznets pointed out that the narrowing inequality should be most likely to occur with the rise in the income share of the poor group in the non-agricultural sector. He gave two explanations for this phenomenon: the first is the self-adaptation effect and the second comes from the support of a democratic institution.

Kuznets (1955) offered a comprehensive argument on the issue that how income inequality is affected during the growth process. Unfortunately, the later studies only focused on the "inverted U" curve but ignored most of the arguments in his paper. In this chapter I will choose to apply some points of Kuznets' arguments on analyzing China's growth-inequality nexus since 1990s. These applications seem very effective.

The later empirical studies have serious controversies on Kuznets "inverted U"

hypothesis. Deininger and Squire (1998) presented a review on relative discussions. They tested the hypothesis with panel data covering 108 countries from 1960s to 1990s. Generally speaking, the Kuznets hypothesis can find more proof from developed countries than developing countries. But from the point of my view, this result just proves the validity of Kuznets hypothesis since the "inverted U" curve describes the process of an economy from developing stage to developed stage. Also, they reported that not all developed countries and developing countries supported Kuznets curve. Methodologically, they found some evidence supporting the hypothesis with cross-section estimation but this result is very sensitive to regional dummies. The estimation with panel data doesn't show any support for the hypothesis. They also tried to test individual countries with time series data. Their findings are not so supportive for Kuznets hypothesis. Some countries' data even show a U-shaped but not "inverted U" curve. In my view, this result mainly comes from the time range of the data. For example, they failed to see any statistically significant relationship between inequality and income for some developed countries and for China. The reason is that for the period from 1960s to 1990s, these developed countries have already stayed at the later industrialization stage and China was a pure planned economy before 1980s.

Barro (2000) also tested the validity of Kuznets "inverted U" curve with a group of panel data covering roughly 100 countries over 30 years since 1960s. He found that the Kuznets curve appeared as an empirical regularity but this relationship couldn't explain most of the variations in inequality across countries or over time. His paper also discussed some individual effects of some factors on income inequality, like international trade, education, etc.

An issue needs to be paid attention is that Kuznets didn't discuss the change of inequality after an economy completed the structural transition. Thus it is not reasonable to doubt the "inverted U" curve with applying data of limited time range from mature developed economies. To understand how growth in these countries affects income inequality is also important and there are some studies focusing on the determinants of income inequality for such economies. We should look at these results

but separate them from the Kuznets hypothesis. Unfortunately, most traditional studies mixed this crucial difference.

Aghion, Caroli, and Garcia-Penalosa (1999) made a review on the discussions on the effect of growth on earnings inequality, esp. the situation in OECD countries from 1970s to 1990s. They agreed with Jan Tinbergen (1975) that it is the acceleration of the growth of the relative demand for skills that increased the skill premium which explains the trend of earnings inequality. They further argued that there are mainly three reasons to explain the acceleration: the impact of trade, the skill-biased technological change and the organizational change within firms. They concluded that economic growth does not necessarily bring a reduction in inequality. For the above three factors influencing inequality, they argued that technological change is the major determinant and channel from growth to inequality since trade effect is carried over by technical change and organizational change is also affected by it. One report of OECD in 2011 (see the Reference) discussed the growing income inequality in OECD countries over the past two decades. The main resources that the greater inequality comes from were identified to be the distribution of wages and salaries which account for 75% household incomes of working-age people. The above three factors continued to be the driving force of such trend plus tax and benefit system change and household structure change.

As mentioned above, checking Kuznets hypothesis needs to cover longer economic development periods and it is different from checking how the inequality is determined for developed countries. Mah (2003) made an empirical research on how the inequality of South Korea was determined from 1975 to 1995. Since South Korea in this period transited from an agricultural economy to a developed economy, he found a support for Kuznets hypothesis and the turning point is between 5000 USD to 6000 USD. But he didn’t find any support for the view that globalization effects (like trade and FDI) influenced the GINI index which was similar to the argument of Aghion, Caroli, and Garcia-Penalosa (1999). This result also supports my view that we should look at the problem of Kuznets hypothesis and income inequality determinants based on development stages. Since China is a transitional economy, we

can suppose that its income inequality change is experiencing part of the Kuznets "inverted U" curve and the structural change will be the main impetus behind it. The later study will initially prove that this hypothesis is true.

The effect of macro policies on income inequality is another topic. Thurow (1970) used Beta distribution to simulate the actual income distribution since he found that the Beta distribution result was fit to U.S. income distribution statistics from 1949 to 1966. He divided the analysis into white people and black people and concluded that growth would be a good policy instrument for more equality in blacks and inflation would be a proper policy instrument for more equality in whites. He also suggested a proper macro policy is needed for less inequality which shouldn't be too vigorous.

There are several different angles in the literatures studying China's income distribution. Zhao, Li and Riskin (1999) argued that the Kuznets curve didn't have the statistic support from China's regional income distribution. Wu and Perloff (2005) analyzed China's income distribution from 1985 to 2001. They showed that the rising income inequality between rural and urban areas in China accounts for most part of the increasing inequality. Wang and Fan (2005) analyzed Chinese regional income inequality from 1980s. They pointed out that the larger regional income inequality mainly came from the increased income inequality among rural income levels in different areas. They argued that more capital lead by market power flew to eastern area of China which caused the greater regional inequality. The factors of human capital and policy design also take effect. Wan, Lu and Chen (2006) discussed the relative problems concerning growth-inequality nexus, using panel data of provinces of China from 1987 to 2001. They failed to find the evidence for the Kuznets curve but found an opposite U curve. Their results showed that the fiscal expenditure on supporting rural development and urbanization significantly reduced the income inequality and more openness increased the inequality. Further, they found that the growth of non-state-owned economy has a negative effect on income distribution. This is also consistent with the argument of Song, Storesletten and Zilibotti (2011) that provinces with more private firms have much higher income inequality.

2.3 An empirical study on explaining China's income inequality

2.3.1 Methodology and data

I apply time series OLS method as the basic estimation method. The method of applying the panel data from provinces of China, in my view, has several problems: firstly, China is a very unbalanced economy. The coastal provinces are much more advanced than the inner land. This means that the cross section or the panel method will possibly bring the problems of measurement errors, heterogeneity and omitted variable bias.

Measurement errors: so far there isn't any official statistics on each province's level of income inequality. Also in China's political institution, the data reported by different local governments are not of the same level of reflecting the truth. The measurement errors could lead to an estimation bias. From the point of my view, there is no way to resolve this problem.

Omitted variable bias and heterogeneity: since different provinces are staying at different development stages, the determinants for the growth may be different. Thus the problem of the omitted variable bias and heterogeneity will easily appear.

The following equation is estimated to examine whether the Kuznets "inverted U" curve appeared in China during the past 30 years:

$$\text{Gini coefficient}(t) = \beta_1 \log(\text{GDP})(t) + \beta_2 \log^2(\text{GDP})(t) + \beta_3 \text{market capitalization}(t) + \beta_4 \text{inflation} + u(t) \quad (1)$$

where u(t) is the conventionally assumed error term. If β_1 is positive and β_2 is negative, it means that the Kuznets hypothesis is valid.

Since China is a transitional economy, based on Kuznets' hypothesis, structural change should be the main impetus behind the change of income inequality. To examine the determinants of income inequality, the following equation is estimated:

$$\text{Gini coefficient}(t) = \beta_1 \log(\text{GDP})(t) + \beta_2 \text{industrial growth rates difference}(t) + \beta_3 \text{government expenditure}(t) + u(t) \quad (2)$$

where u(t) is the conventionally assumed error term and the crucial factor: Industrial growth rates difference =

$$\frac{(\text{growth rates of primary industry}+\text{growth rates of secondary industry})}{2}$$

growth rates of the tertiary industry.

Here tertiary industry is just another expression of service industry. GDP is reported in constant prices with unit of Billions RMB Yuan; Government expenditure is reported as percentage of GDP; Market capitalization is reported as percentage of GDP of the publicly listed companies.

In equation (1) and (2) I include market capitalization, inflation and government expenditure as the additional variables which influence the Gini coefficient. The factor of market capitalization has been rarely applied in such regressions but its influence on income inequality should be clear as a progress of financial market. Inflation and government expenditure are usually treated as negatively influencing inequality and positively influencing inequality respectively.

Data

There isn't any Chinese official Gini index available for the past 30 years and the existing calculations on Chinese Gini index are rather diversified. Measurement error is not a problem that can be avoided and there will be different estimation results if applying Gini coefficients from different studies. Based on a critical attitude, I chose the Gini index calculated by Cheng (2007) as the main sample for estimating income inequality of China. His study results have been frequently discussed and introduced in literatures. Gini coefficients reported in Cheng's study are from 1981 to 2004, except the coefficient for 1991. For the consistency of data standard, I left the Gini coefficient for 1991 for blank. Gini indexes from 2005 to 2010 are chosen from three resources: Gini coefficients for 2005 and 2006 are chosen from Chen and Dai (2011). Gini coefficients for 2007, 2008 and 2009 are reported by The World Factbook of Central Intelligence Agency, USA. It is a common sense that the Chinese Gini index in 2010 is above 0.5, according to a report made by Xinhua Agency published on May 21st, *China Economic Information Daily*. We set it as 0.5. The data are reported in Table 2.1.

Table 2.1

Year	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Gini		0.3993	0.4183	0.43	0.4169	0.3946	0.3964	0.4001	0.4124	0.4275
Year	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Gini	0.4331	0.4297	0.443	0.4419	0.4573	0.4624	0.415	0.469	0.48	0.5

The basic attitude of choosing the above data for Gini index is based on the fact that during the past 5 years Chinese people have been feeling the worsening trend of social justice and equality. The Gini coefficients in Table 2.1, although come from different sources with different calculation processes, can describe China's income inequality change approximately.

Meanwhile, although in January 2013 the Chinese government (National Bureau of Statistics) published the Gini index they calculated (Table 2.2) from 2003 to 2012 with the equality situation getting better and better since 2008, their calculation without issuing the calculation details is highly doubtful. But in order to see whether the empirical result is robust with changing the Gini data, I also performed regressions with changing the data from 2003 into the ones issued by the National Bureau of Statistics, P. R. China. The results are reported in Table 2.8.

Table 2.2

Year	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Gini	0.479	0.473	0.485	0.487	0.484	0.491	0.49	0.481	0.477	0.474

Source: National Bureau of Statistics, P.R. China

Data for the industrial growth rates are counted from the data provided by Annual Statistics Reports, National Bureau of Statistics, P. R. China (Table 1.4). Other data for the period 1991-2010 comes from the *World Bank website Open Database for China* and the website of International Monetary Fund, *World Economic Outlook*

Database, April 2012.

2.3.2 Empirical results

The main regression results reported below (Table 2.3 to Table 2.7) are based on the Gini index in Table 2.1. We start by checking the data stationarity by unit root tests. The results are presented in Table 2.3 with the method of Augmented-Dickey-Fuller (ADF) unit root tests.

Table 2.3

Results from ADF Unit Roots Test for the data from 1991(1992) to 2010

Series	Prob.	Observations
Gini coefficients	0.7869	19
Market capitalization	0.2432	20
Log (GDP)	0.9834	20
Inflation	0.3674	20
Government expenditure	0.9214	20
Industrial growth rates difference	0.0454	20

Note: Null hypothesis is unit root. Lags' numbers are 0 in the ADF equation necessary to eliminate AR errors.

As shown above, only the variable "industrial growth rates difference" is stationary. But a deeper observation shows that three other variables seemingly non-stationary are trend stationary:

Series	Prob.	Lag length
Log (GDP)	0.0718	4
Government expenditure	0.02	3

Market capitalization	0.01	4

Note: Results from ADF Unit Roots Test with a constant and linear trend.

And the rest two variables are non-stationary but stationary with 1st difference:

Series	**Prob.**	**Lag length**
Gini coefficients	0.0011	0
Inflation	0.0293	0

Note: Results from ADF Unit Roots Test with 1st difference.

Since the regressions' models are OLS regression, non-stationary and trend stationary are not crucial for OLS regression assumptions (Wooldridge: *Introductory Econometrics*, pp. 363, pp. 382) except the potential problem of spurious regression. We can resolve the worry of spurious regressions here with checking the cointegration relationship. The unit root tests on the two equations' residuals prove that the residuals are stationary (not reported). Thus the non-stationary variables are cointegrated and the problem of spurious regression is avoided.

The results of regression (1) and (2) are presented in Table 2.4.

Table 2.4

OLS estimation results with time series data from 1992 to 2010

Variable (1)	**Coefficient**	**Variable (2)**	**Coefficient**
Log (GDP)	0.025 **	Log (GDP)	0.041***
	(0.009)		(0.003)
Log^2(GDP)	0.003 **	Industrial growth rates difference	0.896***
	(0.001)		(0.225)
Market capitalization	-0.000 ***	Government Expenditure	0.005***
	(0.000)		(0.002)

Inflation	0.002***	
	(0.001)	
Adjusted R^2	0.81	0.81
No. Obs.	19	19

Note: The dependent variables for both regressions are the Gini coefficients. ** means statistically significant at 5% level.*** means statistically significant at 1% level. Values within the parentheses below the estimated coefficients denote the standard errors.

Some additional tests are also shown to support the results (Table 2.5): Breusch-Pagan-Godfrey tests are performed to test heteroskedasticity. The results show that this problem doesn't exist in these two regressions. The Ramsey RESET tests are performed for these OLS regressions and the results show that the functional forms are appropriate. Residual Breusch-Godfrey LM tests and Durbin-Watson tests are checked for the autocorrelation problem. The results show that our regressions don't suffer from such a problem.

The OLS regression requires the error term obeying the Gaussian distribution (Normal Distribution). We performed the Jarque-Bera tests. The results accepted the null-hypothesis that the residuals obeying the Gaussian distribution. From the skewness and kurtosis values, we can see that regression (2) is better than regression (1) in this standard since samples from a normal distribution have an expected skewness of 0 and an expected kurtosis of 3. According to Jarque and Bera (1987), the critical values of Jarque-Bera test for 20 observations are 2.13 for 10% significance level and 3.26 for 5% significance level, using 10000 replications.

Table 2.5

	Equation (1)	**Equation (2)**
Breusch-Pagan-Godfrey	1.29	0.57
(F-Statistics)		
Ramsey RESET Test	1.47	0.55

(F-Statistics)		
Breusch-Godfrey LM tests	0.10	0.47
(Prob.:Chi-Square (2))		
Durbin-Watson Statistics	1.46	2.13
Jarque-Bera Test	1.46	2.38
(P-Value)	0.48	0.30
(Skewness)	-0.14	-0.86
(Kurtosis)	1.67	3.12

Note: Chi-Square (2) for Breusch-Godfrey LM tests means that there is no serial correlation up to 2 lagged periods.

There are many OLS regressions suffer from the problem of endogeneity or the OLS estimations will not be consistent. For the regressions here, considering that there are many other variables that could explain some part of the Gini coefficients besides the explanatory variables here, the "other variables" may also be correlated with the existing explanatory variables. Further, the level of income inequality may also affect economic performance and thus causes the endogeneity problem.

I apply instrumental variables to resolve the problem in the above two regressions. Since the autocorrelation problem is proved to be non-existed and the regressors in the OLS structure are contemporaneous time series, I choose the lagged variables of the regressors as instrumental variables. Similar choices are taken by Barro (2000). In order to make the test more trustable, I choose two kinds of instrumental variables: the first group takes one-period lagged values as the instrumental variables and the second group takes two-periods lagged values the instrumental variables.

The results from 2SLS regressions show that after adding such instrumental variables, the coefficients of the second regression testing the effect of industrial structural change are still significant (see Table 2.6). The estimation result from adding instrumental variables for the first regression testing the Kuznets hypothesis, however, becomes non-significant (see Table 2.7).

Table 2.6

2SLS estimation results with instrumental variables for regression (2)

Variable	Coefficient 1	Coefficient 2
Log (GDP)	0.041***	0.042***
	(0.004)	(0.004)
Industrial growth rates difference	1.175***	0.903***
	(0.400)	(0.270)
Government Expenditure	0.005**	0.004**
	(0.002)	(0.002)
Durbin-Watson	2.12	2.10
No. Obs.	19	18

Note: The dependent variable is Gini index. Time series data is from 1992 to 2010. The instrumental variables for the column of "Coefficient 1" are one-period lagged value Log(GDP)(-1), Industrial growth rates difference (-1) and Government Expenditure (-1); The instrumental variables for the column of "Coefficient 2" are two-periods lagged value Log(GDP)(-2), Industrial growth rates difference (-2) and Government Expenditure (-2); ** means statistically significant at 5% level. *** means statistically significant at 1% level. Values within the parentheses below the estimated coefficients denote the standard errors.

Table 2.7

2SLS estimation results with instrumental variables for regression (1)

Variable	Coefficient 1	Coefficient 2
Log (GDP)	0.128	-0.035
	(0.192)	(0.102)

Log^2(GDP)	-0.01	0.01
	(0.024)	(0.013)
Market capitalization	0.002	-0.001
	(0.004)	(0.002)
Inflation	0.002	0.003
	(0.004)	(0.002)
Durbin-Watson	1.99	2.17
No. Obs.	19	18

Note: The dependent variable is Gini index. Time series data is from 1992 to 2010. The instrumental variables for the column of "Coefficient 1" are one-period lagged value for all regressors; The instrumental variables for the column of "Coefficient 2" are two-periods lagged value for all regressors. Values within the parentheses below the estimated coefficients denote the standard errors. None of the results are significant.

Testing Over-identification Restrictions

Taking the lagged value as instruments in time series regressions can be seen in many literatures. But considering that the appropriateness of the instrumental variables are always doubtful, I also tested the relationship between the instrumental variables and error term using over-identification restrictions test for regression (2).

The test is divided into three steps: (1) obtain the 2SLS residuals ũ with all the instrumental variables; (2) then regress the residuals ũ on all the instruments to obtain the R-squared value; (3) under the null hypothesis that all instrumental variables are uncorrelated with the original error term u(t), $nR^2 \sim$ Chi square distribution and check it with 5% critical value in the distribution.

The R^2 we get is 0.1 with 18 observations. The 0.05 critical value for the Chi-square distribution with 3 degree of freedom is 7.815. Thus from the result achieved by the above steps we accept the null hypothesis which means that all the instrumental variables are uncorrelated with the original error term.

Robust Test

For seeing whether the empirical results of regressions (1) and (2) are robust with changing the Gini data, I also performed regressions with only changing the Gini data from 2003 into the data issued by the National Bureau of Statistics, P. R. China. The results are reported in Table 2.8.

Table 2.8

Variable (1)	Coefficient	Variable (2)	Coefficient
Log (GDP)	0.031 ***	Log (GDP)	0.041***
	(0.007)		(0.004)
Log^2(GDP)	0.002 **	Industrial growth rates difference	0.629**
	(0.001)		(0.274)
Inflation	0.002***	Government Expenditure	0.005**
	(0.001)		(0.002)
Adjusted R^2	0.84		0.82
Durbin-Watson	1.11		0.92
No. Obs.	19		19

Note: The dependent variables for both regressions are the Gini coefficients with changing the Gini coefficients from 2003 to 2010 into the data reported by the National Bureau of Statistics, P. R. China. ** means statistically significant at 5% level. *** means statistically significant at 1% level. Values within the parentheses below the estimated coefficients denote the standard errors.

Discussions

Based on limited observations, the empirical results initially (with instrumental variables estimation results not significant) proved that the Kuznets “inverted U” hypothesis didn’t appear in China from 1992 to 2010. This result is also supported by the regression result reported in Table 2.8. In regression (1), after changing the Gini

coefficients from 2003 to 2010 into the data reported by China National Bureau of Statistics, the regressor "Market capitalization" is not significant anymore but the other results are generally similar. To be noted, equation (2) is not a proper structure for identifying Kuznets inverted U-shape hypothesis since the regressor $\log^2$(GDP) is not significant in that structure. From the results of estimating equation (1), it seems that China is still staying at the first half the "inverted U" curve that the income inequality is still increasing as GDP grows. Market capitalization and inflation don't show much influence on Gini index. But from the regression on equation (2), surprisingly, the industrial growth rates difference nearly has 90% of the explaining power on both regressions with and without instrumental variables. Although this doesn't mean that 90% of the variation in income inequality is caused by this factor, it means that when industrial growth rates difference increase by 1 unit, Gini coefficient in China will decrease by nearly 0.9 units. Government expenditure, although usually thought as important, didn't play a big role in China. Another 4% comes from GDP growth. Clearly, for the 20 years, higher GDP achievement has a positive effect on the increment of Gini index. I also tried to change the formula of the regressor "industrial growth rates difference" to the other possibilities like differences between the secondary industry and the tertiary industry and between the average growth rate of the sum of secondary and tertiary industries and the primary industry. The results are still significant but either of two factors' influence will reduce to less than 50% (not reported). So the factor of industrial growth rates difference I used here has the biggest explaining power to see the economic structural effect on this issue. Table 8 shows the explaining power of the industrial growth rates difference is robust. After changing the Gini coefficients from 2003 to 2010 into the data reported by the China National Bureau of Statistics, the results are still significant with just lowering the coefficient of the industrial growth rates difference to 63%, which still shows the major role of the factor in explaining inequality.

Although the factor of industrial growth rates difference proves to be closely related to income inequality, very few literatures paid attention to this as reported in the session of "Literature Review". The result also proves the argument of Kuznets (1955)

that the industrial structure update has a major role in shaping the trend of income inequality. The difference between the empirical results and the Kuznets' argument lies in the tertiary industry's role. Kuznets (1955) divided the industrial structure into agricultural and non-agricultural which means he didn't talk about the specific role played by the tertiary industry. In our estimation structure, the tertiary industry plays a determinant role in affecting industrial growth rates difference since the larger the growth rate of tertiary industry, the smaller the difference will be. This also means that the role and effect of tertiary industry on the change of income inequality should be paid more attention. Although the other direction that reducing the sum of the growth rates of the primary and secondary industries can also reduce the degree of industrial growth rate difference, this is not the case we should consider since it is the direction opposite to growth promotion.

2.4 The role of tertiary industry and growth-inequality nexus

The global economic history shows that the economic structural update is accompanied by continuous economic growth and meanwhile the Kuznets "inverted U" curve appears during the whole transitional process. As a style fact, the developed economy experiences the process from primary industry domination to tertiary (service) industry domination as a share of GDP. So the emergence of tertiary industry is usually seen as the natural result of economic growth. Also, the empirical studies in this chapter show that the effect of industrial update of tertiary industry has a very significant explaining power on income inequality in China. The higher the growth rate of tertiary industry, the lower the income inequality will be. Thus the Kuznets' arguments on the "inverted U" curve are also supported by the empirical results in this chapter, that is, the income inequality will naturally lower down as the economic growth continues to the stage of tertiarisation.

However, our empirical results also generated a more complicated picture. This chapter shows that from regression (1), we didn't see the "inverted U" curve happened in China yet, with continuously higher growth rates of the secondary

industry than those of the tertiary industry (See Figure 1.5). In my view, China has long ago passed the point that the tertiary industry should have grown faster than the secondary industry but its development has been blocked by the economic and political institutions which will be explained in the next section. The neoclassical growth economics, however, generally assumes that tertiarisation process will be automatically realized during growth and they don't pay any attention to the possibility of institutional block for successful tertiarisation. But this chapter's results also support the attitude of neoclassical economics towards the evolution of income inequality during growth. Differently, I would argue that economic growth is also the result of successful industrial update from primary to tertiary and this process can be stopped or blocked by many reasons.

Let's further discuss the economic role of tertiarisation from both the demand and supply sides. For the demand side, the tertiarisation process is usually explained by the higher income elasticity of the tertiary industry since people will consume more in tertiary industry with rising income (Chenery, 1982). The rising income, however, comes from continuous economic growth without very high inequality in income/wealth distribution. If the income inequality is very serious, the demand power will be deterred even with high growth rates since richer people have a higher propensity to save (Kaldor, 1955). Stiglitz (2012) also presented the similar argument that the income inequality will lead to insufficient demand for supply growth. We should also be aware of that demand from outside the country cannot support the growth of tertiary industry since most of the service is non-traded products. China is exactly an example that its secondary industry has been promoted by exports for a long time and its tertiary industry cannot find sufficient support from trading. For this reason, the income inequality generates a negative effect on tertiarisation. China's case shows that tertiarisation could be very hard to complete under very high economic growth rate. We should be aware of that the very high economic growth rates in China were supported during early period of economic development which doesn't mean that they will be supported during the later stage when tertiarisation should exert more influence. Further, our empirical results show that the effect of

industrial update of tertiary industry also has a very significant explaining power on income inequality in China. This means that the tertiarisation process further impact the income/wealth distribution in an economy. This seems there are mutual effects between tertiarisation and income/wealth distribution. The effect of tertiarisation on income inequality can be explained from the supply side.

For the supply side, the tertiarisation is the result of continuous economic growth which means that the sustainable technology innovation in the economy is the crucial factor behind tertiarisation. Tertiarisation is the process of enhancing overall productivity since it is updating the resource to higher value-added industries. And this is the reason why the average wage rate in the tertiary industry is higher than that of the other two industries. The tertiary industry has a higher value-added per labor or capita input, meaning the production resource will be used more efficiently. After the mature stage of secondary industry when its marginal profit rate is suffering the law of diminishing marginal returns, a competitive economy will enter the stage of tertiarisation automatically and the tertiarisation process will also increase the productivity of both secondary and primary industries. This is the reason for the higher average wage rate which lowers the inequality degree. It should be noted that the effect of tertiarisation on inequality is different from that of technology innovation.

Some literatures also mentioned that the industrial transitional process is partly endogenous to entrepreneurial response (See Peneder, 2002). Remember both Karl Marx and Joseph Schumpeter emphasized the entrepreneurs' role on continuous technology innovation. This logic also points to the importance of an institution which supports entrepreneurial spirit.

Traditionally, tertiarisation is seen as a natural result of economic growth and its effect on growth is treated as much smaller than within industry productivity growth. Peneder (2002) empirically tested the effect of industrial update on economic growth and the result showed that the effect was not as big as expected since most growth power came from productivity progress within each industry. However, he didn't check how much the productivity progress within each industry is affected by

tertiarisation. These previous studies on the effect of industrial update on growth, from the point of my view, didn't fully understand the effect of tertiarisation on growth. Particularly, these understandings haven't clearly seen the importance of successful tertiarisation on continuous economic growth for transitional developing economies. The above discussion on the role of tertiarisation shows that there is an inner link among growth, distribution and tertiarisation: tertiarisation is the result of continuous economic growth but it also means that only with successful tertiarisation the economic growth can be sustained from a developing to a developed economy. The reason is that for the demand side tertiarisation is the condition of continuously sufficient demand with a more equal distribution to support the supply growth and for the supply side tertiarisation is the realization process of higher overall productivity. McMillan and Rodrik (2011) empirically confirmed that most difference of growth rates between developing countries was contributed by update of overall labor productivity which is determined by structural change. The above logic also leads to the crucial effect of equal income/wealth distribution on long-run economic growth.

To sum up, the above argument points to the importance of growth with equal distribution for a transitional developing economy. The belief that unequal distribution is good for growth (esp. in the long-run) is strongly challenged here. So what kind of economy can realize the sustainable growth from a developing stage to the developed stage which means that the tertiarisation process will be successful? For understanding the answer to this question, it is a good idea for us to see what can block the progress of tertiarisation and China is the right case for this phenomenon.

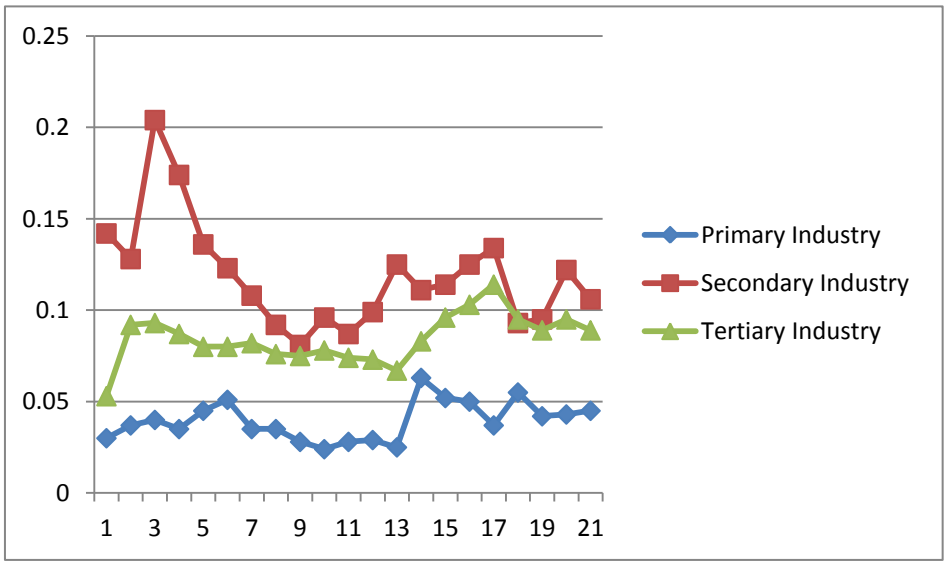
2.5 Why tertiary industry develops slowly in China? An institutional analysis

The economic structure of China has experienced a big change during the past 20 years. In 1990, 73.6% of the population was still living in rural areas producing 27.1% products of GDP. These numbers updated to 27.1% for rural population and 11.3% for GDP contribution in 2009 (Lin, 2011b). The growth rates of secondary and tertiary

industries are also very high.

From Figure 1.5, we can see that the growth rates of tertiary industry in China during the past 20 years are generally lower than the growth rates of secondary industry. From the view of employment rate, it is reported that by the end of 2011, the employment share of tertiary industry reached 35.7% which is bigger than that of secondary industry (29.5%) and primary industry (34.8%) (*China People's Daily*, 2012.6.5). But the state sector still keeps 77% of total urban employment in tertiary industry in 2007 which contrasts 15% in manufacturing (Zhu, 2012). From the perspective of output, according to China National Bureau of Statistics, in 2011 the tertiary industry's contribution to GDP is 43.1% which is much less than that of developed economies where the ratio is usually larger than 70%. This ratio is even lower than the countries in the similar development stage like India. Considering the bubble in China's real estate market and real estate is a big part of tertiary industry, the tertiary industry is even smaller in China if the bubble effect is omitted.

Figure 1.5 Industrial output growth rates between 1991 and 2011, China



Note: The column line shows the output growth rates and the row line shows the time range.

Source: Data collected from the Annual Statistics Reports, National Bureau of Statistics, P. R. China.

Also, the international growth experience of the developed economies shows that the

growth rate of tertiary industry will be higher than that of secondary industry when an economy enters the stage of tertiarisation. Obviously, we haven't seen China reached this stage. However, China has already reached the stage of high middle-income level according to World Bank and should have already entered the period of tertiarisation. As argued above, tertiarisation is the condition for generating sufficient demand to support growth. But China's growth experience has already proved that the process is rather difficult to realize, e.g. the 2004 annual Central Economic Work Conference of China decided to alter China's growth path to rely more on the domestic consumption and the central government have tried many ways to realize the goal. So far, however, the goal is still far from reach. Then what is blocking China's steps toward a tertiary industry-dominated economy? This is the question we should try to answer.

The "three" crucial institutional arrangements

I would like to argue that the basic reason for slow tertiarisation in China is deeply rooted in its institutional arrangements. The marketing economy in China now is far from a real competitive market. To the opposite, China is a very high rent-seeking economy. From the point of my view, there are three crucial institutional arrangements to be primarily responsible.

Firstly, China's political institution makes this country a very high rent-seeking economy. There are three influential contents of the political institution. The first part is the difficulty to realize the high-level law-ruled society under the one-party leadership, although China has made much progress to a law-ruled society since 1979. Under current political institutions, Chinese government officials still own much power to intervene the judiciary process since the judiciary sector in China is not independent. Further, their power on managing public resources is weakly limited and the public monitoring power in China is very weak. As a result, the corruption phenomenon is quite prevalent. This makes real fair competition very hard to realize.

The second part is that China's working units are divided into "outside institution" units and "inside institution" units. The "inside institution" means the part controlled by CPC and the promotion channel has the same ranking arrangements. The "inside

institution" units include all levels of governments, some kinds of enterprises (like the State-Owned Enterprises) and public-sector units. These units have a natural closer relationship with political power and their rent-seeking space is much larger than the "outside institution" units. Especially, nearly all the financial sectors in China are "inside institution" units which keep the monopoly status.

The third part is the institutional arrangements on how the local officials' performances being measured and judged. In fact, the promotion of the local officials is mainly based on two kinds of reasons: the political relationship with the higher-level officials and the economic performance (GDP) of the place in charge. This institution is called "local competition" which is argued by some economists like Steven Cheung to be the most important growth motivation behind China's growth miracle. For this incentive arrangement, every local government chief is trying the best to build the relationship with senior officials and to expand the GDP of the place they stay. The former one leads more corruptions and the later one leads to the investment-led growth model since the investment is the quickest way of increasing GDP growth rate.

Besides, China has a long tradition of "*Guan Xi*" culture. For thousands of years, China is a country ruled by people but not ruled by law or any religion. Also, the traditional Chinese culture represented by Confucianism emphasizes the effect of clan. With such complicated historical reasons, the Chinese "*Guan Xi*" which means relationship has been a kind of culture prominent in China. Fairness and justice are always influenced by "*Guan Xi*". Because of the "*Guan Xi*" culture as well as the political institution which supports the culture to take effect, China is a country with very large rent-seeking space. The people and organizations with closer and better "*Guan Xi*" have larger rent-seeking space. Under many situations, the rent-seeking activities can exert the determinant effect on varieties of issues.

Secondly, China is adopting a double-track economic system. The National Development and Reform Commission of China has the authority to decide prices for many resources and products, like agricultural products, water, oil, steel, electricity, medicine, transportation, etc. These resources are mainly controlled by State-Owned

Enterprises (SOEs) with monopoly power. This is a long tradition formed in pure planned economy after 1949 that governments can keep the input prices low enough to subsidize industrialization. Although many sectors are open to private economies, the most important economic sectors in the economy are still owned by State-Owned Enterprises (SOE). Although the reform is marketing economy oriented, the CPC still holds a belief that the so-called "socialism marketing economy" should be labeled by the existence of SOEs which takes the main part in an economy.

With this economic structure, the price system of the double-track economic system has two levels: the first level is that although there are official prices for some commodities determined by the central government, there are also market prices for these commodities, like agricultural good. The second level is that the prices for some goods (e.g. the oil) are strictly determined by the central government and the prices for some goods are purely determined by the market. Thus the whole price system of the double-track economy is complicatedly distorted. Further, the SOEs are mainly capital-intensive which are supported by very low interest loans from state-owned banks and their relationships with governments are very close. In China, the Chairman of the Board of SOEs has the same rank as the officials in the governments. The ranks are between the levels of Director and Minister. To the opposite, the private economies are very difficult to get loans since most loans go to SOEs. Wu (2000) reported that state sector controls much of China's capital. Kuijs and Wang (2005) reported that the small and medium-sized companies which were mainly private received less than 10% to total bank loans. Most of the private companies are thus labor-intensive and they feel very hard to update into capital-intensive industries.

Thirdly, the economy is a dual-sector economy dividing into rural and urban sectors. The system of household registration ("*Hukou*"), although has been loosening, is still the main block for labor flows. The dual economy makes a very unequal development between rural and urban areas which limits the income growth of rural habitants. Most of them don't have many welfare supports like limits on children's education and verities of insurances. Only 20% of migrants get access to urban health insurance in 2010 and 12-13.5% of migrants get access to unemployment insurance from 2008 to

2010 (Meng, 2012). Another crucial economic result is that the wage rate of migrant laborers in cities is thus depressed with discriminations and occupational choices. Rural laborers in China are always the weaker side in contract games, compared with the rights of urban citizens. Their hourly wage rate on average is only 45% of the average urban workers who own the "*Hukou*" (Meng, 2012).

The effect of the institutional arrangements

The above three institutional arrangements strongly distorted China's market and affected both the demand and supply sides for the successful tertiarisation.

Firstly, the private economy has to face the competition of SOEs with very unfair conditions. This leads the Chinese economy to a very strange combination: the private economy is forced to develop labor-intensive industries which are mainly in the secondary industry (China National Bureau of Statistics, 2004). They firstly entered the low-end tertiary industry in the early reform period as the owner of some small businesses like tiny restaurants and education service which have smaller scale compared to those in primary and secondary industries. However, tertiary industry is marked with both capital intensive and labor intensive. The industries of finance, transportation and real estate, for example, are important parts of tertiary industry. Most of these capital-intensive and very high-value added industries are occupied by SOEs. Lin and Li (2001) also argued that in China most of the capital-intensive enterprises are SOEs. The private economies are very difficult to compete with them without equal treatment by institutions. Many tertiary capital-intensive industries are still monopolized with entry limitation. Most (70%, Lin (2011a, Chapter 9)) of the commercial loans of "inside institution" financial sectors are flowing to SOEs.

However, the SOEs entering the tertiary industry have very low efficiency. Lin and Tan (1999) developed a concept called "viability" for analyzing SOEs in a socialism economy. Lin (2011a) explained this concept as:

"Viability" is the capacity of a normally managed enterprise to earn a socially acceptable normal profit in an open, free and competitive market, without external

support or protection. "Normally managed" means that there is no major problem in its operation or management. "Normal profit" refers to an average profit acceptable to the market. "Open" means that the domestic market is connected with foreign markets. "Free" means free access to the market. "Competitive" means that there is no monopoly in the market.

With this concept, he argued that the China SOEs existing with varieties of governmental supports and privileges are actually non-viable and endogenously bring the distorted banking system and the regulation of market access.

Liu (2000) found that the SOEs have the lowest efficiency in China, using input-output method. Lu (2003) reported that the efficiency difference between SOEs and private economy had been keeping increasing, although the efficiency of SOEs also somehow improved. There are many literatures explaining the phenomenon which can be summarized as huge rent-seeking space enjoyed by SOEs. One argument is about the "soft budget constraint" which means that the SOEs don't have real budget constraint as private economies and thus their incentive to improve the efficiency is very low. Further, their investments with the loans are generating huge waste which also causes a big burden for the banking system. The private sectors are thus difficult to get sufficient and timely loans for development. This leads to long-time monopolies in many fields and lack of competition. Besides, SOEs (and the state-related sectors which are "inside institution") in China are low efficient but they can offer higher wages and welfare to the employees. This causes a high misallocation of both human resource and physical capital: the best talents compete to enter the SOEs but their high marginal productivity is limited by SOEs' institution. Meanwhile, the competition is not fair among the talents who are trying to enter the state-related sectors (including the SOEs) because of the rent-seeking activities which lead a mismatch between the high-payment (or welfare) jobs and the talents who own the relative qualified productivity. The system of SOEs also supported governmental interventions into the economy. Kuijs and Wang (2005) argued that more investment will enter industries other than the secondary industry and labor resource can also be

allocated more efficiently without governmental interventions. Since the capital-intensive industries in tertiary industry are occupies by such SOEs, the growth rate of tertiary industry is highly depressed. To the opposite, the secondary industry grows faster because private economies are forced to concentrate on labor-intensive manufacturing industries and luckily, China indeed has a comparative advantage of labor-intensive industries compared to developed countries. This is reason behind the picture why the growth rate of tertiary industry has been keeping lower than the secondary industry. However, the comparative advantage of such private enterprises in the secondary industry has been diminishing as Chinese economy grows and this leads to the structural problem more serious.

Further, SOEs have very low incentive to update their productivity as a response to market competition because they don't have much competition pressure. The lack of entrepreneurial spirit of SOEs blocks the efficiency update including technology innovation. Meanwhile, the private companies which are very sensitive to market competitions are lack of capital to burden the cost of technological innovation. They also have much weaker support from government and thus are in shortage of many other "rents" compared to SOEs. There is indeed a competition between the private economy and state-owned economy in China. However, the competition is not fair which greatly reduces the competition pressure on SOEs. And since the competition is closely related to "rent", most private companies are also focusing on "*Guan Xi*" building (investment) but not productivity promotion. The competition for rent-seeking further makes the rents more prevalent and more expensive which lead the economy more unfair. There are some models arguing the efficiency of the double-track economic system, e.g. Lau, Qian and Roland (1997, 2000) modeled the double-track reform strategy and argued that it is not only efficient but also Pareto-improving. Their models, however, are far from reality. There isn't "rent" in their models and the planned economic units work as the private economic units. Thus these models do not fit China's reality during reform. Murphy, Shleifer and Vishny (1993) argued that there are two reasons why rent-seeking is very costly to growth. The first reason is that rent-seeking activities exhibit very natural increasing returns

which attract more and more resource to rent-seeking activities from productive activities. The second reason lies in that rent-seeking, esp. public rent-seeking by governmental officials, is prone to hurt innovative activities since innovations often need public supports like patterns and licenses. Both these two reasons obviously exist in China whose economic and political institutions limit the update speed of technology. The above arguments revealed the reasons why China is a fast-growing economy but lacks internationally competitive large companies. Also, the average wage rate of the whole economy is highly depressed and the domestic demand is thus insufficient.

As argued above, the double-track economic institution as well as political institution blocked the natural growth of Chinese private economy. The urban-rural dual system which limits the national labor market also negatively influences the tertiarisation process. We can see this from several angles:

firstly, the house registration system blocked the shape of a competitive labor market which distorted the labor price. The migrant laborers from rural area stay at a weaker situation in gaming wage contract. An even worse effect is that rural migrants are nearly impossible to enter SOEs to enjoy the higher income with welfare and most of them have to enter private economies with accepting very low wage rate. Only 7.3% of rural migrants are employed in the state sector in 2009 (Meng, 2012). To summarize, the depressed low labor cost supports and stimulates the private economy in labor-intensive industries, esp. in the manufacturing secondary industry.

Secondly, since there is much welfare loss of rural laborers caused by the house registration system, the migration process from rural to urban areas is blocked. The amount of money they can transfer back to their rural families is also limited. From Figure 1.5 we see that the growth rates of agricultural industry have been much lower than the other two industries. One important reason is the slow and even repeated migration process of rural laborers since the quick and sustainable migration is one of the key conditions for raising productivity in agriculture during economic transition. The above logic means that the household registration system limited the income growth of around more than 60% of Chinese people who are still living in rural areas

and who moved to urban areas without "*Hukou*". This also means that the national average education level is limited by the system. So from both the demand side and the supply side, the registration system restricts the conditions required for continuous tertiarisation.

The above arguments put forward the basic reasons behind the slow development of tertiarisation in China. Behind the high GDP growth rates, China is a highly distorted economy: all the product market, resource market and labor market are distorted by such institutional arrangements. The whole economy is thus low efficient and weak. Hsieh and Klenow (2009) estimated the effect of China's resource misallocation on TFP and they calculated that China's TFP will increase around 30%-50% if capital and labor in China are assumed to equalize marginal products to the extent like that in United States. Many studies already pointed out that the raising of TFP or efficiency is the only way that China can sustain its high growth rates and realize its growth potentials and this needs basic institutional reforms.

2.6 Policy suggestions

As a suggestion for China's future reform, the successful and quick tertiarisation is the goal that should be focused. The chapter argues that China's tertiarisation is slow because of three main institutional arrangements: political institution, double-track economic system and rural-urban dual economy. Accordingly, the relative future reforms should focus on these institutions and this will also be a process of deprivation of rents.

To be more exact, the current institutional reforms should focus on the privatization for the main industrial update blocks, esp. liberalizing the tertiary industry; abolishing the system of household registration and building a high-quality law-ruled society. These reforms are connected with each other and should be promoted at the same time or the high growth rate of China will face unsustainable problem in the near future and the traditional problem of "middle-income trap" will appear. Although Chinese government has already paid attention to develop tertiary industry, there is one trap of

tertiarisation China should avoid: this chapter doesn't mean that continuing tertiarisation with more state-owned economy can also be a choice. The tertiarisation process must be completed with economic liberalization, or many of the arguments in section 2.4 cannot be established any more. The Chinese government, therefore, should not only look at the development speed of tertiary industry but also pay attention to the development path. This seems different from the arguments of Wan, Lu and Chen (2006) and Song, Storesletten and Zilibotti (2011) that the private economy exerts a negative influence on China's income/wealth distribution. Their arguments are not wrong, however, from the static view. But from a dynamic viewpoint, continuing economic liberalization should be right choice.

2.7 Pro-poor growth? A discussion with Dollar and Kraay (2002)

China's reform and opening up during the past 30 years has made a great achievement in reducing absolute poverty. According to China National Bureau of Statistics, the number of people living in absolute poverty reduced from 300 million to less than 100 million which occupied the biggest share of global poverty reduction. But whether growth is a pro-poor process, based on the concept of relative poverty (a concept related to distribution, sometimes defined the bottom 10% or 20% of income distribution as the poor group), has been debating by economists. A recent most well-known result was achieved by Dollar and Kraay (2002). In the paper *Growth is Good for the Poor*, they made an empirical study based on 92 countries spanning the past four decades. Dollar and Kraay defined the "poor" as those in the bottom fifth of the income distribution of a country. Their regressions found that the general relationship between growth of the income of the poor and growth of mean (average) income is nearly one to one. Besides, their result is very strong since they also found that many other factors traditionally treated as influencing factors actually don't matter at all for this relationship like the development periods and international trade.

According to their results, we can have several implications:

Firstly, the starting point of the poor in different countries matters. If the starting

income of the poor in two countries are different and their GDP grow rates keep the same, then the poor people will never have the same income in the two countries. Secondly, considering the real world, we can see that countries have different GINI index during different development stages. For example, in early 1980s, China's GINI index was very low (less than 0.2) but now it is high (more than 0.5). According to the result of Dollar and Kraay (2002), if the per capita of China grows 9% since 1980s, the poor people's income has also increased at nearly the same speed. The result of Dollar and Kraay (2002) actually changed our thought on the "Lorenz Curve": suppose that in 1980, the bottom fifth people have 20% of the total income in China, thus in 2010 the bottom fifth people will also have 20% of the total income in China! The point (0.2, 0.2) in the Lorenz Curve will keep the same. We know that the Lorenz Curve of China in the past 30 years has greatly lowered the right downward. So the main reason to explain the change of GINI index will be that most of the change comes from the change of the other 80% people. One big problem lies in the work of Dollar and Kraay (2002) is that they didn't show the result for each percentage below 20%, thus we don't know the change of points in Lorenz Curve below (0.2, 0.2). But since it is relatively much smaller, we can see that much of change is generated by the other 80% people.

The above argument leaves many important questions which Dollar and Kraay didn't explain:

1. What is the economic theory to explain such a result that the income change of the rich people (as defined 80% of the population) should be a major contribution to the GINI index change?

2. The result shows that the policy direction of achieving a good GINI index should mainly take care of the other 80% population under economic growth. Then what types of policies should be the optimal choices?

Of course the above argument and questions rise only if the work of Dollar and Kraay is solid. However, Lubker, Smith and Weeks (2002) gave a fierce criticism on it. They put forward that the work of Dollar and Kraay has several serious flaws:

1. The empirical work is based on theoretically unsound equations;

2. The data are seriously flawed;

3. The policy variables are not defined appropriately.

This paper can be seen as a comprehensive criticism. Besides, if their comment is solid, then the regression result of Dollar and Kraay will be very doubtful. Dollar and Kraay in the same year quickly gave a defending on their work and pointed out the criticisms of Lubker, Smith and Weeks are incorrect. Their response, however, didn't cover all the problems pointed out by the three economists. Ravallion (2001) argued that both the pro-poor growth argued by Dollar and Kraay (2002) and the opposite position of anti-poor growth have some supports.

Since Dollar and Kraay (2002) established their results based on panel data covering 92 countries, it is meaningful and interesting to see whether a single transitional economy like China obeys the rule they set up.

Table 2.9 Income share held by lowest 20% people in China

Year	1993	1996	1999	2002	2005
Share	7.35%	7.24%	6.39%	5.47%	4.99%

Source: World Bank Website Open Database for China

The above data shows that the income share held by the lowest 20% people in China decreased by 32% during the 12 years. We also have the Gini coefficients were 0.4183 in 1993 and 0.4573 in 2005 which means that Gini coefficients increased 9.3% during the 12 years. These data shows several implications: the first is that China's growth didn't seem to follow the rule of Dollar and Kraay (2002) that growth is relatively good for poor with the relationship between growth of income of the poor and growth of mean income is one-to-one, although the absolute poverty problem has been greatly improved in China during the reform period. Also the change of "Lorenz Curve" cannot be explained only by the richer 80% population. The result of Dollar and Kraay (2002) is thus strongly challenged by China's experience. But if we define

the “pro-poor” growth as alleviating poverty, the answer will be “yes”: growth is pro-poor.

2.8 Concluding Remarks

This chapter empirically tested the Kuznets “inverted U” curve with economic data of China since 1990s. During this period China successfully transited from a poor economy to a high middle-income level economy. The estimation results initially show that China seems still at the first half of the “inverted U” curve.

The study shows that the industrial growth rates difference can explain most part of the income inequality change in China. The role of tertiarisation is very crucial for the change. We also discussed the importance of tertiarisation on sustainable growth. Finally, we argued that the development of tertiary industry is slow in China because of three main institutional arrangements which are political institution, double-track economic system and rural-urban dual economy.

The study also confirms a link between growth and distribution that is the low income inequality should be good for growth in the long run, especially during the later stage of development when tertiarisation should become dominated. A proper economic institutional arrangement should be good for sustainable growth and equal income/wealth distribution in the long run as well. To the opposite, the worsening income inequality which is accompanying slow tertiarisation makes the sustainable growth difficult.

Chapter 3

Inequality Evolution with Rent-seeking: from the Specific Case of China to Generalization

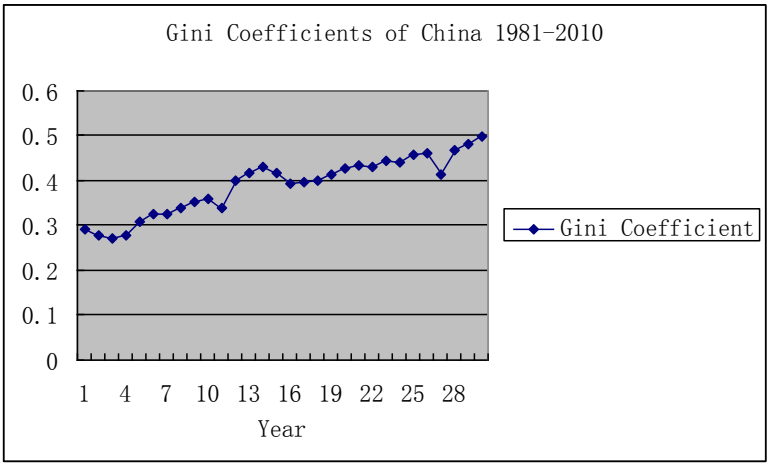
3.1 Introduction

It is worldwide acknowledged that China has been a fast-growing economy with worsening income/wealth distribution for over 30 years since 1979. Figure 1.1 and Figure 1.2 depict the details of both growth and distribution performance of China since 1980s (we use income Gini index as the proxy for wealth Gini index). The average growth rate reaches about 10% and the Gini coefficient has been keeping increasing. How to under this phenomenon is an interesting question and under hot debate.

Figure 1.1

Note: Growth rates from 1980 to 1990 are calculated by GDP given by Maddison (2007) and growth rates from 1991 to 2010 are calculated by GDP given by International Monetary Fund, World Economic Outlook Database, April 2012.

Figure 1.2



Note: Gini coefficients from 1981-2004 (except 1991) are the results calculated by Cheng (2007); Gini coefficient for 1991 is the result calculated by Wang (2010); Gini indexes from 2005 to 2010 are chosen from several resources: Gini coefficients for 2005 and 2006 are chosen from Chen and Dai (2011). Gini indexes for 2007, 2008 and 2009 are reported by The World Factbook of Central Intelligence Agency, USA. It is a common sense that the Chinese Gini index in 2010 is above 0.5, according to a report made by Xinhua Agency published in May 21st, *China Economic Information Daily*. We set it as 0.5.

There are many perspectives on how to explain the growing inequality in China. Lin and Liu (2003) argued that the increasing regional inequality of China rooted in the development strategy which doesn't obey the comparative advantage. Wang and Fan (2004) analyzed Chinese regional income inequality from 1980s. They pointed out that the larger regional income inequality mainly came from the increased income inequality among rural income levels in different areas. They argued that more capital lead by market power flew to eastern area of China which caused the greater regional inequality. The factors of human capital and policy design also take effect. Wang (2007) emphasized the role of grey revenue in explaining China's income inequality.

I would like to look at the problem from a rent-seeking perspective. Let's see who became rich in China. According to Gerth (2010), China's newly rich class after reform and opening up can be divided into four groups: the individual entrepreneurs ("*getihu*"), people who took advantage of the double-track price system, people who are land speculators and managers of state-owned enterprises (SOEs). I would like to point out that there is another group obviously exists: the governmental officials of all

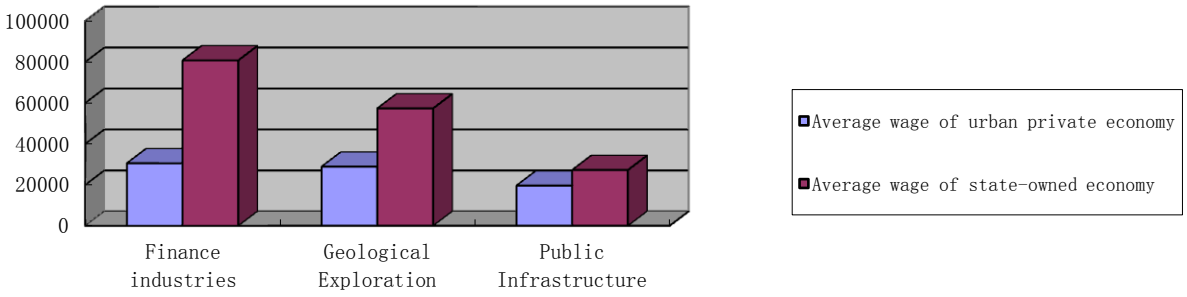
levels. And the poorest people in China are mainly the poorest in rural areas, most "off-post" workers of SOEs and most rural migrants in cities. It should be noted that the newly rich class, as introduced, has a close relationship with rent-seeking.

The "rent" we mentioned here is neither the Ricardian land rent nor the quasi-rent we usually mentioned in economics. According to Tollison (1982), rent-seeking is "*the expenditure of scarce resources to capture an artificially created transfer*" which is pointing to the activities of wasting resources competing for artificially contrived transfers. Rent, which often being treated as related to government powers, can be interpreted as a kind of privilege. These pursuing activities for rent can be competitive. However, the actions of rent-seeking cause pure waste and reduce the efficiency of competition for the overall society. Early analyses were mainly static which pointed out the welfare loss caused by rent-seeking activities of monopoly actions or international trade (see Tullock, 1967 and Krueger, 1974). Murphy, Shleifer and Vishny (1993) put forward two reasons why rent-seeking is very costly to economic growth. The first reason is that rent-seeking activities exhibit very natural increasing returns which attract more and more resource to rent-seeking activities from productive activities. The second reason lies in that rent-seeking, esp. public rent-seeking by governmental officials, is prone to hurt innovative activities since innovations often need public supports like patterns and licenses.

China is an economy with very large rent-seeking space which is shaped by its institutions. Under current Chinese political and economic institutions, the economy can be divided into two parts: inside institution ("*tizhinei*") and outside institution ("*tizhiwai*"). This is the result of China's gradual reform since 1979. The central government adopts the way of loosening the planned economy step by step. The groups inside the institution mean the ones controlled or at least partly controlled by the governments and they enjoy much more political power compared to the outside institution economy. The working units inside the institution contain varieties of governmental organizations and many enterprises. Thus the scope of this group is very large. For example, most Chinese universities are public and all of them are belonging to the group inside the institution. A very realistic incentive for the people

working in such units is that they have a possibility to be promoted according to the official rank (same as the governmental officials) whereas the people working in the units outside the institution don't. As a result, the units inside the institution enjoy much larger rent-seeking space than the groups outside the institution. A well-known example in China is that the employees inside the institution can buy the real estate through their unit welfare system which will enable them to buy the house with much lower price (e.g. 10% of the market price).

Figure 1.6 Average wage difference between sectors in 2010



Source: Reports of the National Bureau of Statistics, China

If we divide China's economy into two industries of competitive private economy which is outside institution and monopolistic state-owned economy which is inside institution, we will find that not only the average wage is much higher in the state-owned industry (see Figure 1.6), but also its wage growth rate is much higher. According to Banister (2007), the average real annual wage growth rate in unban manufacturing sector during 1992 to 2004 was 7.5% but if excluding state-owned and collectively owned economy, the growth rate becomes 4.6%. This is a unique phenomenon because the state-owned economy is lack of efficiency. There are large amount of studies showing that the productivity of Chinese private economy is much higher than that of the state-owned economy.

The above distorted economic phenomenon is shaped by Chinese institutional arrangements with "Chinese characteristics". In this chapter, I try to model and explain the phenomenon through including the institutional factors into the

neoclassical economic models. Observing that the economy is divided into two parts, we built Model 1 and 2 employing two representative agents. Each represents a position to the governments. The economic analysis on Chinese economy cannot ignore the effect of rent which means that we need to revise and extend the traditional neoclassical growth models. The models we built in the chapter show that the groups within the institution and the ones outside it have two different convergence points. This means the income inequality will be permanent without institutional innovation and the Kuznets "inverted U" curve will not appear. The models also show that transferring the "double-track" economy to a "single-track" one will boost the whole economy. The chapter finally generalized the models to show that the equal distribution of rent-seeking space will lead to equal wealth distribution dynamically. Here the rent-inequality nexus has been established. The chapter is to reveal and analyze the deep reason behind the Chinese Gini index performance but not to quantitatively explain the Gini index determinants.

The content is arranged as follows: section 3.2 explains the incentives for rent-seeking; section 3.3 studies the wealth distribution evolution under neoclassical growth models with two representative agents; section 3.4 creates a generalized model to establish the rent-inequality nexus; section 3.5 is the concluding remarks.

3.2 Why the agents will perform rent-seeking?

In an economy constructed by "inside institution" enterprises and "outside institution" enterprises, a simple game analysis will show that both the two kinds of enterprises will perform rent-seeking which is a "Nash Equilibrium" solution.

We assume the existing institution is fixed. Such an institution determines the payments of the game. Following the fact that rent-seeking can bring additional revenue and the agents inside institution own higher rents, the payments for the "outside institution" economy (represented by private enterprises) and the "inside institution" economy (represented by SOEs) are shown in Table 3.1:

Table 3.1

		SOEs	
		Rent-seeking	Rent-seeking Deny
Private Economy	Rent-seeking	4, 6	10, 0
	Rent-seeking Deny	3, 10	9, 5

SOEs in China are acknowledged efficiency-lack. Liu (2000) found that the SOEs have the lowest efficiency in China, using input-output method. Lu (2003) reported that the efficiency difference between SOEs and private economy had been keeping increasing. There are many literatures explaining the phenomenon. One argument is about the "soft budget constraint" which means that the SOEs don't have real budget constraint as private economies. This is a result of rent-seeking. Besides governmental subsidies, most (70%, Lin (2011, Chapter 9)) of the commercial loans of "inside institution" financial sectors are flowing to SOEs. With the "soft budget constraint", their incentive to improve the efficiency is very low. Further, their investments with the loans are generating huge waste which also causes a big burden for the banking system. The private sectors are thus difficult to get sufficient and timely loans for development. This leads to long-time monopolies in many fields with huge rent-seeking space and lack of competition.

Based on this, we assume that the highest profit can be earned is 10 units for both of the two types of enterprises after considering cost. If both the SOEs and private enterprises don't perform rent-seeking activities, the economy will stay at the fair competition status. Since the SOEs are much less efficient than the private economy, the private economy will earn the profit 9 units and the SOEs will only get 5. If only the SOEs perform rent-seeking activities, they will get 10 units and the private economy will get 3. This is the result under extreme unfairness. The private economy can still get 3 units because of their high efficiency. If only the private economy perform rent-seeking, the SOEs will get nothing and the private economy get 10.

Finally, if both the agents perform rent-seeking activities, neither of them can get 10 units and SOEs can stay at an advantage with getting 6 units and the private economy get 4 units.

These payments assumption is based on economic logic which can be adjusted. For example, we may assume that the profit of SOEs is greater than 0 when they don't pursue rent-seeking and the private economy performs rent-seeking. But the profit cannot surpass that under the rent-seeking condition.

The above payments show that there is a Nash equilibrium at the situation that both agents perform rent-seeking. The total payment will be lower than that under fair competition. This reflects the negative externality of rent-seeking activities. The institutional arrangements affected the revenue and cost of each agent's activity. Thus it should be noted that the institutional innovation will change the Nash equilibrium through changing the payments.

One thing needs to be noted is that the Nash equilibrium above is not a Prisoner's Dilemma. Rama (1993) argued that rent-seeking can be understood as a repeated Prisoner's Dilemma. This argument isn't consistent with Chinese economic reality. The reason is that SOEs and the private enterprises have different efficiencies. If they compete at an environment without rent, the profit gained by SOEs will be definitely lower than that gained with rent, although the total profits will become larger.

3.3 How income/wealth distribution evolves during economic growth with rent?

3.3.1 Literature review

The earliest neoclassical model on the evolution process of income/wealth inequality during economic growth can be traced to Stiglitz (1969). Stiglitz developed an analysis based on the Solow growth frame. He assumed a linear savings function which is totally exogenous, a constant depreciation rate, homogeneous labor and equal wealth division among one's heirs. In such a growth frame, he managed to prove that the wealth and income will be asymptotically evenly distributed, whatever

the initially wealth distribution is. This result is also solid to nonlinear saving functions. But his study also shows that if the labor force is heterogeneous, the wealth distribution will be determined by the productivity distribution. Stiglitz (2012) reviewed this research in the book *The Price of Inequality: How Today's Divided Society Endangers Our Future* and pointed out that the asymptotic even distribution is just a dream in the real world. He discussed the role of rent in the new book. His research in 1969, however, induced many researches later. Bourguignon (1981) deepened the analysis of Stiglitz (1969), proving that with the convex saving function, there can be two steady states with one equal distribution and the other one unequal distribution, depending on the inequality of the initial income distribution. Interestingly, his model showed that the unequal stationary distributions are "pareto superior" to the equal one, that is, his arguments tend to treat the unequal distribution as a good matter.

From 1990s, most of the researches gave up the assumption of exogenous saving function and adopted the Ramsey optimization frame. Lucas (1992) discussed the distribution problem in a dynamic economy under different allocation mechanisms. The results are diversified and there is no unique trend of distribution evolution. Chatterjee (1994) did the pioneering work on the distribution problem under a standard one-sector neoclassical growth model. In his work, the individual wealth growth is related to the savings which is determined by optimization. This is a Ramsey-Cass-Koopmans frame which differs the Stiglitz (1969)'s methodology. He showed that with some parameter restrictions, the wealth inequality will become higher when the economy stays at the situation of more than the steady-state per capita capital stock, vice versa. This is similar to the Kuznets inverted U-curve argument. Thus the economy will converge to a steady state of inequality. But this result can also be reversed if change some of the parameter restrictions. His paper shows that the degree of inequality will lower down without government intervention and stabilize at a lower level in the long run.

Casselli and Ventura (2000) tried to apply the representative consumer theory to analyze the problem of income/wealth inequality in the neoclassical growth models.

They argued that the representative consumer theory doesn't rule out consumer heterogeneity. They also argued that the representative agent model can be seen as having the average characteristics of an economy and thus they are able to construct the relative values of a_i/a, c_i/c and y_i/y where the note i is the individual identify. My critics in the later section will show that the above arguments are not solid. They further analyzed the dynamic evolution of the relative values in the Ramsey-Cass-Koopmans model and the Arrow-Romer model. The results of wealth distribution dynamics are diversified with different setups. The choice of Cobb-Douglas technology or the CES technology affects the results for the Ramsey-Cass-Koopmans model whereas in the Arrow-Romer model the results will be even more diversified by choosing different specifications.

Li, Xie and Zou (2000) also analyzed dynamics of wealth distribution within extended Ramsey-Cass-Koopmans model and endogenous growth model. The methodology they used is different from Casselli and Ventura (2000). They hope to get the evolution trend of individual wealth as aggregate capital stock goes larger. The results are that under the Ramsey frame, the inequality will converge to the steady-state capital stock; under the endogenous frame, the inequality will continue to expand or stays the same, depending on the setup of production function. The weak point of their paper is that the assumptions made during the calculations are not flexible and purely mathematical technique without well-argued economic intuition. They also made empirical tests with a panel data set to find the supports for the Ramsey-Cass-Koopmans frame and doubt the validity of the endogenous frame.

3.3.2 Neoclassical models with two representative agents

The first model to be built here is an extended Ramsey-Cass-Koopmans model. We hope to check how the inequality evolves under the assumption of continuous technological progress. Thus an exogenous frame is sufficient. The traditional Ramsey-Cass-Koopmans model doesn't assume the existence of any economic rent. It also excludes the problem of distribution with only one representative agent.

A Critic on the usage of one "representative consumer" on the problem of

income/wealth distribution

It has been a tradition to apply "one representative agent" to analyze the problem of income distribution for some years. Caselli and Ventura's paper (2000) has been treated as the pioneering work on this approach. They argued that the "representative consumer" doesn't rule out consumer heterogeneity and can be treated as having the average value characteristics. They argued that the Gorman form requirement doesn't affect the application of this theory into distribution problem. These viewpoints, however, are not well-grounded and they did not argue in detail in their paper either.

Traditionally, economists modeling the economic environment with one representative agent by explicitly assuming that all the agents are homogenous. This is well-grounded for the following reason: the Gorman form requirement for the indirect utility function of each individual's utility function implicitly assumes that the wealth distribution is equal for all the agents.

The Gorman form requirement for consumer n can be described as:

$v^n(p,w)=a^n(p)+b(p)w^n$, where the w^n is consumer n's initial wealth, p is the price level.

It is proved that if all consumers have indirect utility functions of the Gorman form, total demand of an economy can be written as a function of prices and total wealth and this is a necessary and sufficient condition (Mas-colell, A., Whinston, M.D. and Green, J. R. (1995)) A further analysis on the problem that whether the distribution of wealth matters for the above result shows that if for a fixed price each consumer's wealth expansion path is a straight line and the slopes of the lines are the same, then wealth distribution won't affect the total demand of an economy. In mathematics, the condition is: $\frac{\partial x_i^n(p,w^n)}{\partial w^n}=\frac{\partial x_i^m(p,w^m)}{\partial w^m}$ where i represents the good i, n and m represent consumers.

The above condition implicitly means that every consumer has the same wealth because only with the same wealth, the marginal change of the consumption of good i can be equal when their wealth changes. The Gorman form thus assumes a world of equality and the form doesn't mean the average amount of an economy in the real

world. As a result, the one representative consumer theory cannot be applied to analyze the distribution problem.

Model 1: Decentralized agents problem

The above critic doesn't rule out the possibility of assuming more than one representative consumer in an economy and every representative consumer represents a group of people whose wealth expansion paths are paralleled. In this section I consider a closed economy with two representative agents: two representative consumers and two representative firms. The design of this is based on the argument that China's economic resources are divided into two parts: "inside institution" and "outside institution" according to the relationship with governments. The firms "inside institution" are mainly SOEs and the firms "outside institution" are private economic agents.

Households

It is a tradition to apply the unit of household as the basic demand side. This setup is especially good for the growth-inequality study. An obvious fact is that the equal distribution should be considered from the perspective of household but not individuals. The education level of a father cannot be changed but this change can be realized by his children. The better wealth situation of the children usually means the parents are also becoming economically better. So the inter-generational change during economic growth is meaningful for social equality with this household setup.

Similar to the setup of Ramsey-Cass-Koopmans model, we consider a continuous-time economy with an infinite number of long-lived agents. We assume that there are two representative households (consumers) in the economy. Household i's instantaneous utility function is u(c(t)) where $u(c(t)) = \frac{c_i^{1-\theta}-1}{1-\theta}$, i=1 and 2 with $\theta > 0$ and $\theta \neq 1$ which is the inverse of the intertemporal elasticity of substitution. We also suppose that the labor growth rate is e^{nt} with L(0)=1. Every household hopes to maximize the total utility U(t) where $U(t)= \int_0^{\infty} e^{-\rho t} e^{nt} u(c(t)) dt$. The

assumption of u(t) is strictly increasing, concave, twice continuously differentiable with derivatives $u'(c)$ and $u''(c)$ as well as satisfies the Inada condition: $\lim_{c\to 0} u'(c) = \infty$ and $\lim_{c\to\infty} u'(c) = 0$. The factor ρ represents the time preference rate and $\rho > n$ which ensures the U(t) is bounded. We assume that both ρ and θ are the same for the two representative households.

Let a_i denote the average amount of asset the individual i holds in the representative household, r_i denote the risk-free market flow rate of return of assets and w_i denote the average individual labor income (wage), we have the familiar evolution equation: $\dot{a_i}(t)=(r_i(t)-n)a_i(t)+w_i(t)-c_i(t)$. This equation applies to both the two representative consumers with $a_i(0)$ given.

The transversality condition is $\lim_{t\to\infty} a(t)\exp\left(-\int_0^t (r(s)-n)ds\right) = 0$. This condition is to ensure the individuals would never want to have positive asset when the final (if there is) period finishes.

Maximize U(t) subject to $\dot{a_i}(t)$ equation can generate the well-known Euler equation for the problem: $\frac{\dot{c_i}(t)}{c_i(t)} = \frac{1}{\theta}(r_i(t)-\rho)$, c=C/L.

Enterprises

Rama (1993) already tried to describe the effect of rent seeking in the representative firm production function. Following this approach, we assume that there exist two types of enterprises. For simplicity, we use the notations i and ĩ to represent the enterprises of the two different types. For the enterprises outside institution, the production function plays as $Y_i=F(K_i, AL_i, R_i, R)=K_i^{\alpha}(AL_i)^{\beta}(\frac{R_i}{R})^{\gamma}$. R_i represents the rent it owns and R represents the average rent of the whole economy.

This setup has several implications: if $\frac{R_i}{R} > 1$, it means that the firm's rent level is greater than the average and thus enjoys the revenue brought by its rent; if $\frac{R_i}{R} < 1$, it means that the firm's rent level is smaller than the average and thus the firm suffers from the negative effect of total rent. We assume that $\alpha+\beta+\gamma = 1$ which means that the firm's technology is neoclassical.

For the enterprises inside institution, the production function plays as $Y_{ĩ}=F(K_{ĩ}, AL_{ĩ}, R_{ĩ})=K_{ĩ}^{\alpha}(AL_{ĩ})^{\beta}(R_{ĩ})^{\gamma}$. For simplicity, we assume the two kinds of enterprises enjoy the same parameters α, β and γ. $R_{ĩ} > 1$ means that all the enterprises inside institution take advantage of its individual rent and don't suffer from the negative externality of the total rent. The variable "A" represents technology and we assume its growth rate is "g" which is totally exogenous.

Discussions

The distribution of rent to the two kinds of economy is based on China's institutional arrangements and current economic reality. The enterprises inside institution doesn't suffer from the negative externality of the total rent of the economy because their monopoly power and privileges are assured by the institution whereas the outside institution economy will suffer from the negative externality of the total rent, although they can also engage in building "*Guanxi*" and enjoy the result of rent-seeking.

The two representative agents

From the above setup, we can get the marginal product of resource K and L for the two kinds of the enterprises:

For the "outside institution" enterprises,

$r_i=\frac{\partial Y_i}{\partial K_i}=\alpha K_i^{\alpha-1}(AL_i)^{\beta}\left(\frac{R_i}{R}\right)^{\gamma}$ and $w_i=\frac{\partial Y_i}{\partial L_i}=K_i^{\alpha}\beta(AL_i)^{\beta-1}A\left(\frac{R_i}{R}\right)^{\gamma}$.

For the "inside institution" enterprises,

$r_{ĩ}=\frac{\partial Y_{ĩ}}{\partial K_{ĩ}}=\alpha K_{ĩ}^{\alpha-1}(AL_{ĩ})^{\beta}R_{ĩ}^{\gamma}$ and $w_{ĩ}=\frac{\partial Y_{ĩ}}{\partial L_{ĩ}}=K_{ĩ}^{\alpha}\beta(AL_{ĩ})^{\beta-1}AR_{ĩ}^{\gamma}$.

We can see the difference between the two representative enterprises of this setup by calculating the per capita income y, denoting y=Y/AL, k=K/AL and c=C/AL. For the "outside institution" enterprises, we have

$$y_i=\frac{K_i^{\alpha}\left(\frac{R_i}{R}\right)^{\gamma}}{(AL_i)^{\alpha+\gamma}}=\frac{K_i^{\alpha}R_i^{\gamma}}{(AL_i)^{\alpha}(AL_iR)^{\gamma}}=k_i^{\alpha}\left(\frac{R_i}{AL_iR}\right)^{\gamma}$$

For the "inside institution" enterprises, we have

$$y_{\tilde{\imath}} = \frac{K_{\tilde{\imath}}^{\alpha} R_{\tilde{\imath}}^{\gamma}}{(AL_{\tilde{\imath}})^{\alpha+\gamma}} = k_{\tilde{\imath}}^{\alpha} \left(\frac{R_{\tilde{\imath}}}{AL_{\tilde{\imath}}}\right)^{\gamma}$$

Since the rent of "inside institution" enterprises are much higher than the other, $R_{\tilde{\imath}} \gg R_i$. Also the employment amount L of "outside institution" enterprises is much larger according to the reality of Chinese economy, we have $L_i \gg L_{\tilde{\imath}}$ (the results is not affected if we assume that $L_i = L_{\tilde{\imath}}$). Suppose that they have the same amount of capital, we will have $y_{\tilde{\imath}} \gg y_i$.

From the expressions of y, we can also get the representative households' revenues:

$$r_i = \alpha k_i^{\alpha-1} \left(\frac{R_i}{AL_i R}\right)^{\gamma}, \quad w_i = k_i^{\alpha} \beta A \left(\frac{R_i}{AL_i R}\right)^{\gamma} = \beta A f(k_i)$$

$$r_{\tilde{\imath}} = \alpha k_{\tilde{\imath}}^{\alpha-1} \left(\frac{R_{\tilde{\imath}}}{AL_{\tilde{\imath}}}\right)^{\gamma}, \quad w_{\tilde{\imath}} = k_{\tilde{\imath}}^{\alpha} \beta A \left(\frac{R_{\tilde{\imath}}}{AL_{\tilde{\imath}}}\right)^{\gamma} = \beta A f(k_{\tilde{\imath}})$$

It is obvious that $r_{\tilde{\imath}} \gg r_i$ and $w_{\tilde{\imath}} \gg w_i$.

Equilibrium analysis:

This is a closed economy divided into two groups with two representative agents. Assume both the whole market and the markets of the two "inside institution" and "outside institution" sectors are complete and clearing, the per capita capital will equal to the per capita assets $k_i=a_i$ for either of the two "inside institution" and "outside institution" sectors where i=1 and 2. When i=1, it means the "outside institution" economy; when i=2, it means the "inside institution" economy, which was represented by $\tilde{\imath}$ in the "enterprises" section. For simplicity, in the chapter we assume the depreciation rate δ is zero. The general equilibrium of the extended Ramsey economy is defined as below:

DEFINITION 1. For either of the two "inside institution" and "outside institution" sectors, the economy stays at an allocation $\left\{\left(k_t^i, c_t^i\right), \left(K_t^i, L_t^i, R_t^i, R_t\right)\right\}_{t=0}^{\infty}$ and a price path $\left\{r_t^i, w_t^i\right\}_{t=0}^{\infty}$, i=1 and 2, such that:

(i) $\left\{r_t^i, w_t^i\right\}_{t=0}^{\infty}$ and $\left\{K_t^i, L_t^i, R_t^i, R_t\right\}_{t=0}^{\infty}$ maximize household's utility and firm profit, for every i;

(ii) The capital and labor markets clear all the time for either of the two sectors, given

initial capital stock K(0) and labor quantity L(0).

Dynamics System

Since we introduced technology progress rate g and normalized y=Y/AL, k=K/AL and c=C/AL, the dynamic conditions will be adjusted to be:

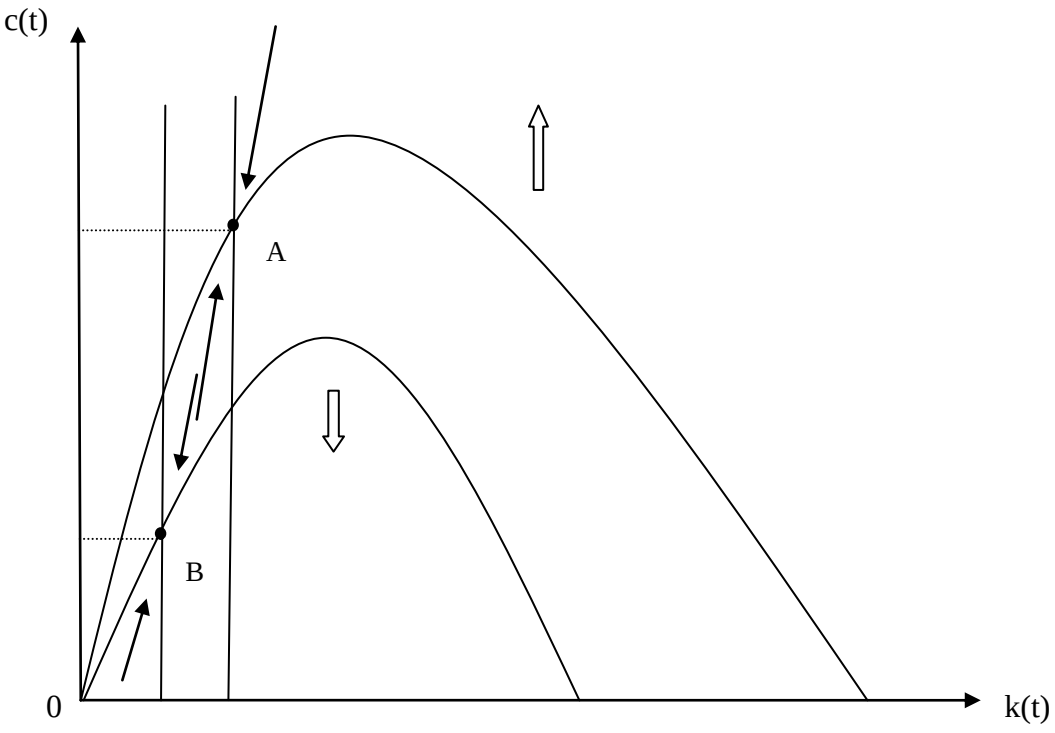
$$\frac{c_i\dot{(t)}}{c_i(t)} = \frac{1}{\theta}(r_i(t) - \rho - \theta g)$$

$$k_i\dot{(t)} = (\alpha + \beta A)f\big(k_i(t)\big) - c_i(t) - (n+g)k_i(t)$$

where i=1 and 2 and $y_i = f\big(k_i(t)\big)$.

With the quantitative relationships analyzed above, we know that the value of $(\alpha + \beta A)$, ρ, θ, g and n are constants for both of two sectors. $r_2 \gg r_1$, $y_2 \gg y_1$. In steady state, we have $\frac{c_i\dot{(t)}}{c_i(t)} = 0$ and $k_i\dot{(t)}$=0. From the quantitative relationships above, we have $c_1^*(t) < c_2^*(t)$ and $k_1^*(t) < k_2^*(t)$, * means the steady state value. These relationships can be described with the diagram in Figure 3.1:

Figure 3.1



The convergence points of the two sectors stay at different places. Point A is the steady state of "inside institution" economy and the point B is the steady state of "outside institution" economy. The phase diagram shows that the consumption level of the two types of consumers will converge to two different points with point A>>point B. There is no power to limit the separation between A and B. This can explain why the Kuznets curve doesn't appear in China. Further, when rent of the "inside institution" economy goes larger, point A will go higher and point B will go down which means that the inequality situation will become worse.

It should be noted that in this kind of setup, even the equal distribution of rent cannot lead to equal income/wealth distribution since only the "outside institution" economy suffers from the total amount of rent. But the depreciation or dilution of rent of the "inside institution" economy will lead the whole economy more equal. This analysis means that this kind of institutional arrangement is extremely bad for the equality of an economy.

Model 2: The Central Planner Problem

The key problem of Model 1 is the assumption that general equilibriums exist in both of "outside institution" and "inside institution" economies. In this section we consider the social planner problem which can avoid such an assumption. Also, Model 1 cannot be applied to analyze the effect of inequality on growth. In this section we consider an economy with endogenous rent-growth through the investment on rent. Assume that the capital is allocated to investment and rent, following the similar laws of motion with the same depreciate rate δ. For simplicity, assume that we don't have the population growth and technological progress here.

There is a benevolent social planner whose goal is to maximize social welfare. There is only one representative consumer whose neoclassical utility function becomes $U(t)=\int_0^{\infty} e^{-\rho t}u(c(t))dt$. The instantaneous utility function is u(c(t)) where $u(c(t)) = \frac{c^{1-\theta}-1}{1-\theta}$, with $\theta > 0$ and $\theta \neq 1$ which is the inverse of the intertemporal elasticity of substitution. There are two representative enterprises whose production functions are

Y_i and $Y_{\tilde{i}}$. The social planner chooses a plan so as to maximize the utility subject to the resource constraint of the economy, taking K(0) and R(0) as given:

Max $U(t)=\int_0^{\infty} e^{-\rho t}u(c(t))dt$

s.t. $\dot{K} = I_K - \delta K$

$\dot{R} = I_R - \delta R$

$Y_i=F(K_i, AL_i, R_i, R)=K_i^{\alpha}(AL_i)^{\beta}(\frac{R_i}{R})^{\gamma}$.

$Y_{\tilde{i}}=F(K_{\tilde{i}}, AL_{\tilde{i}}, R_{\tilde{i}})=K_{\tilde{i}}^{\alpha}(AL_{\tilde{i}})^{\beta}(R_{\tilde{i}})^{\gamma}$.

$Y= Y_i+ Y_{\tilde{i}}=C+I=C+I_K+I_R$

where $R_i+R_{\tilde{i}}=R$.

K(0) and R(0) are given.

We can apply the Hamiltonian method to solve this problem. For simplicity, we assume that $\alpha + \beta = 1$ and $\gamma = 1$. This assumption makes the individual production functions: Y_i and $Y_{\tilde{i}}$ keep the neoclassical characteristics when we move the rent factor to the Y side. We further assume that the "outside institution" and "inside institution" sectors have the same K and L. We define the R here as the sum of the rents which is different from the definition in Model 1.

Define $x = \frac{R_i}{R}$, we have

$Y= Y_i+ Y_{\tilde{i}}=A^{\beta}L^{\beta}K^{\alpha}\left[R_i^{\gamma}R^{-\gamma} + (R - R_i)^{\gamma}\right] = A^{\beta}L^{\beta}K^{\alpha}[x + (1 - x)R]$ (1)

The Hamiltonian problem is:

$J = e^{-\rho t}u(C) + \mu(I_K - \delta K) + \lambda(I_R - \delta R) + \nu\{A^{\beta}L^{\beta}K^{\alpha}[x + (1 - x)R] - C - I_K - I_R\}$

from which we can deduce the first order conditions:

$\frac{\partial J}{\partial C} = 0$ (2) $\frac{\partial J}{\partial I_K} = 0$ (3) $\frac{\partial J}{\partial I_R} = 0$ (4)

together with conditions: $\frac{\partial J}{\partial K} = -\dot{\mu}$ (5) and $\frac{\partial J}{\partial R} = -\dot{\lambda}$ (6)

we can obtain the Euler equation which tells us the growth rate of consumption:

$\frac{\dot{C}}{C} = \frac{1}{\theta}\{\alpha A^{\beta}L^{\beta}K^{\alpha-1}[x + (1 - x)R] - \delta - \rho\}$ (7)

The part of $\alpha A^{\beta}L^{\beta}K^{\alpha-1}[x+(1-x)R]$ is exactly the meaning of interest rate which is r_i in the above section. Since we don't have the technological progress speed g here, the structure (7) is exactly the same as what we get in the Ramsey model.

From (5) and (6) we can get the relationship between K and R:

$$K=\frac{\alpha[x+(1-x)R]}{1-x} \quad (8)$$

Substitute the expression of K in equation (1) by the expression in (8), we have:

$$Y=A^{\beta}L^{\beta}\frac{\alpha^{\alpha}[x+(1-x)R]^{\alpha+1}}{(1-x)^{\alpha}} \quad (9)$$

From (9) we can see that when $x\to 1,\ Y\to+\infty$. This means that all the rent should be owned by the "outside institution" economy in order to maximize the total output. This result is consistent with the fact that the "outside institution" economy is far more efficient than the "inside institution" economy. As an implication, it also means that the "double-track" economy should be transferred to the "single-track" with abolishing the "inside institution" economy.

Discussion

This model is very similar to the classical one-sector growth model with physical and human capital. But the construction of the production functions made our model different. The problem of this model is that one of the implications of this model (also Model 1) is the monotone increasing relationship between the total rent R and the production quantity Y. This could come from three reasons: firstly, the model cannot reflect the opportunity cost made by the rent-seeking activities; secondly, the model doesn't assume the negative influence of rent on economic efficiency, that is, the parameters α and β are independent of R; thirdly, the assumption that the inside institution economy doesn't suffer from the negative externality of the total rent is very strict. We just use this setup to show the different institutional effects for the two kinds of economies.

It should be noted that this central planner problem is set up with negative externalities which means the planner doesn't resolve the externality problem under optimization process. This is consistent with Chinese economic reality that the

government tries to optimize allocation conditional on the existing institutional arrangements.

3.4 Generalization of the models: the rent-inequality nexus

In this section we study the general relationship between rent and wealth distribution within a dynamic frame. We give up the setup of the two representative agents. The basic setup is a revision of the Ramsey-Cass-Koopmans model. One specialty here is that we assume the rent R_i in an economy obeys a certain distribution Ω where $i \in \{1, 2, \cdots, n\}$ representing the heterogeneous households.

Households

We normalize the number of each household population as 1 which means that we don't consider the labor factor L here. Household i's instantaneous utility function is u(c(t)) where $u(c(t)) = \frac{c_i^{1-\theta}-1}{1-\theta}$, $\theta > 0$ and $\theta \neq 1$ which is the inverse of the intertemporal elasticity of substitution. Every household hopes to maximize the total utility U where $U=\int_0^{\infty} e^{-\rho t}u(c(t))dt$. The assumption of u(c(t)) is strictly increasing, concave and satisfies the Inada condition: $\lim_{c\to 0} u'(c) = \infty$ and $\lim_{c\to\infty} u'(c) = 0$. The factor ρ represents the time preference rate which is greater than 0.

Let a_i denote the amount of asset the household (or individual) i holds and r_i denote the risk-free market flow rate of return of assets. Since our main aim is to study the rent-inequality nexus, for simplicity, we omit the wage factor in this modeling like the classical AK model (Rama (1993) also dropped the variable of wage). Thus we have the familiar evolution equation: $\dot{a_i}(t)=r_i(t)a_i(t)-c_i(t)$, with $a_i(0)$ given and is supposed to be 1. Since we assume r_i is not a function of time later, the evolution equation comes into use is $\dot{a_i}(t)=r_ia_i(t)-c_i(t)$. The transversality condition also changes to $\lim_{t\to\infty} a(t)\exp(-rt) = 0$.

Maximize U(t) subject to $\dot{a_i}(t)$ equation can generate the Euler equation for the problem:

$\frac{\dot{c_i}(t)}{c_i(t)} = \frac{1}{\theta}(r_i - \rho)$. We also assume $c_i(0)=1$.

Enterprises

We assume the production function of the enterprises is $Y_i = K_i^{\alpha}\left(\frac{R_i}{R}\right)^{\gamma}$. K here represents capital, R_i represents individual rent and R represents the average rent of the economy. We don't consider the factor "AL" here for simplicity. The production function is similar to the structure of the "outside institution" economy in the above sections. We don't have the "inside institution" economy here in order to make the model general. However, the factor $\frac{R_i}{R}$ is still there since rent-seeking activities exist in all kinds of the economies.

The marginal product of capital thus becomes $r_i=\frac{\partial Y_i}{\partial K_i}=\alpha K_i^{\alpha-1}\left(\frac{R_i}{R}\right)^{\gamma}$. Since we don't assume that the factor of labor L plays a role in the production function, the wage rate is omitted. So this is an economy without considering the labor revenue. Dynamically the wealth inequality is just determined by the asset return which is related to rent.

Equilibrium Analysis

The general equilibrium of the economy is defined as below:

DEFINITION 2. The competitive equilibrium of the economy with heterogeneous agents is an allocation $\left\{\left(k_t^i, c_t^i\right), \left(K_t^i, R_t^i, R_t\right)\right\}_{t=0}^{\infty}$ and a price path $\left\{r_t^i\right\}_{t=0}^{\infty}$, $i \in \{1, 2, \cdots, n\}$, such that:

(i) $\left\{r_t^i\right\}_{t=0}^{\infty}$ and $\left\{K_t^i, R_t^i, R_t\right\}_{t=0}^{\infty}$ maximize household's utility and firm profit;

(ii) All markets clear all the time.

From the asset evolution equation we have

$$\dot{a_i}(t) = \alpha K_i^{\alpha-1}\left(\frac{R_i}{R}\right)^{\gamma} a_i(t) - c_i(t)$$

From $\frac{\dot{c_i}(t)}{c_i(t)} = \frac{1}{\theta}(r_i - \rho)$ we have

$$c_i(t) = c_i(0)\exp\left\{\frac{1}{\theta}\left[\alpha K_i^{\alpha-1}\left(\frac{R_i}{R}\right)^{\gamma} - \rho\right]t\right\}$$

We assume that in the above equations only a_i and c_i are functions of time. The other variables and parameters are independent of time. Except for the mathematical reason that such setup is much simpler for calculation, these assumptions can help us analyze the dynamic path of $a_i(t)$ distribution related to certain rent distribution as explained later.

Thus we can get the expression of $a_i(t)$:

$$a_i(t) = a_i(0)\exp\left[\alpha K_i^{\alpha-1}\left(\frac{R_i}{R}\right)^{\gamma}t\right] - c_i(0)\int_0^t e^{\frac{s}{\theta}\left[\alpha K_i^{\alpha-1}\left(\frac{R_i}{R}\right)^{\gamma}-\rho\right]} e^{\int_s^t \alpha K_i^{\alpha-1}\left(\frac{R_i}{R}\right)^{\gamma} ds} ds$$

$$= e^{r_i t} - \int_0^t e^{\frac{s}{\theta}(r_i-\rho)+r_i(t-s)} ds$$

$$= e^{r_i t} - e^{r_i t}\int_0^t e^{\left(\frac{r_i}{\theta}-r_i-\frac{\rho}{\theta}\right)s} ds$$

$$= e^{r_i t} - e^{r_i t}\frac{1}{\left(\frac{r_i-\rho}{\theta}-r_i\right)}\left[e^{\left(\frac{r_i-\rho}{\theta}-r_i\right)t} - 1\right] \qquad (1)$$

$$= e^{r_i t}\left(1 - \frac{1}{\left(\frac{r_i-\rho}{\theta}-r_i\right)}e^{\left(\frac{r_i-\rho}{\theta}-r_i\right)t} + \frac{1}{\left(\frac{r_i-\rho}{\theta}-r_i\right)}\right)$$

$$= \frac{1}{-\left(\frac{r_i}{\theta}-\frac{\rho}{\theta}-r_i\right)}e^{\left(\frac{r_i}{\theta}-\frac{\rho}{\theta}\right)t} + e^{r_i t}\left[1 - \frac{1}{-\left(\frac{r_i}{\theta}-\frac{\rho}{\theta}-r_i\right)}\right] \qquad (2)$$

In order to make sure $a_i(t) \geq 0$, from (1) we know that the condition is

$$\frac{1}{\left(\frac{r_i-\rho}{\theta}-r_i\right)}\left[e^{\left(\frac{r_i-\rho}{\theta}-r_i\right)t} - 1\right] \leq 1 \qquad (3)$$

From (3) we have the boundedness condition[1]:

$$\frac{r_i}{\theta} - \frac{\rho}{\theta} - r_i \leq -1.$$

From (2) we can check the relationship between $a_i(t)$ and r_i. If $\frac{r_i}{\theta} - \frac{\rho}{\theta} - r_i = -1$, $a_i(t) = e^{(r_i-1)t}$. It is obvious that $a_i(t)$ is monotone increasing with respect to r_i. When $\frac{r_i}{\theta} - \frac{\rho}{\theta} - r_i < -1$, we can calculate the first order derivative of $a_i(t)$ with respect to r_i and then we have:

$$\frac{\partial a_i(t)}{\partial r_i} = e^{-\frac{\rho}{\theta}t}\frac{t}{\theta}\frac{1}{\left(1-\frac{1}{\theta}\right)r_i+\frac{\rho}{\theta}}e^{\frac{r_i}{\theta}t} + te^{r_i t} - \frac{\left(1-\frac{1}{\theta}\right)}{\left[\left(1-\frac{1}{\theta}\right)r_i+\frac{\rho}{\theta}\right]^2}e^{\frac{r_i}{\theta}t-\frac{\rho}{\theta}t}$$

$$+ te^{r_i t}\frac{1}{\left(\frac{1}{\theta}-1\right)r_i-\frac{\rho}{\theta}} - e^{r_i t}\frac{\frac{1}{\theta}-1}{\left[\left(\frac{1}{\theta}-1\right)r_i-\frac{\rho}{\theta}\right]^2}$$

We can see that if t is large enough, the above equation is positive[2]. Thus $a_i(t)$ is monotone increasing with respect to r_i when t is large enough.

Then we have

PROPOSITION 1. In an economy with the above setups, when the rent distribution is more equal, the wealth distribution will be also more equal as time goes to be sufficiently long.

Proof: we assume the rent R_i in an economy obeys a certain distribution Ω where $i \in \{1, 2, \cdots, n\}$. R is the average rent in the economy and $\frac{R_i}{R}$ represents the status of rent owned by individual i relative to the average rent. Caselli and Ventura (2000) firstly applied the concept of variance as the measure of inequality. We follow this method to explore the inequality-rent nexus. Define $x = \frac{R_i}{R}$ and we hope to see the relationship between its variance and the variance of $a_i(t)$ based on the equation (2) when time t goes to infinity. It is obvious that r_i is monotone increasing with respect to x. So $a_i(t)$ has the same monotone increasing relationship with x too. Then we assume one case for continuous distribution and one case for discrete distribution in order to make the result more robust.

Firstly, we assume that the distribution of x is the uniform distribution:

It is obvious that $Var[a_i(t,x)]=0$ for any t>0;

Secondly, we assume that the distribution of x is the two-point distribution with equal probability 1/2: suppose that the average rent of the economy is 1, one group of rents stays at 1/2 and the other stays at 3/2. With such a rent distribution, $E[a_i(t,x)] =$

$\frac{1}{2}a_i\left(t,\frac{1}{2}\right)+\frac{1}{2}a_i\left(t,\frac{3}{2}\right)$, and thus $\mathrm{Var}[a_i(t,x)]=\frac{1}{2}\left[\frac{1}{2}a_i\left(t,\frac{1}{2}\right)-\frac{1}{2}a_i\left(t,\frac{3}{2}\right)\right]^2+\frac{1}{2}\left[\frac{1}{2}a_i\left(t,\frac{3}{2}\right)-\frac{1}{2}a_i\left(t,\frac{1}{2}\right)\right]^2=\frac{1}{4}\left[a_i\left(t,\frac{3}{2}\right)-a_i\left(t,\frac{1}{2}\right)\right]^2$. Then we change the setup as one group of rents stays at 1/3 and the other stays at 5/3. This means that the rent distribution is more unequal with variance of the first setup as 1/4 and the second 4/9. With the second rent distribution, $\mathrm{E}[a_i(t,x)]=\frac{1}{2}a_i\left(t,\frac{1}{3}\right)+\frac{1}{2}a_i\left(t,\frac{5}{3}\right)$ and $\mathrm{Var}[a_i(t,x)]=\frac{1}{4}\left[a_i\left(t,\frac{1}{3}\right)-a_i\left(t,\frac{5}{3}\right)\right]^2$. Since $a_i\left(t,\frac{5}{3}\right)-a_i\left(t,\frac{1}{3}\right)>a_i\left(t,\frac{3}{2}\right)-a_i\left(t,\frac{1}{2}\right)$, $\mathrm{Var}[a_i(t,x)]$ with the second setup is also higher. This means that more equal rent distribution will lead to more equal wealth distribution.

Q.E.D.

Discussions

The distributions assumed above are the simplest ones in both continuous and discrete distributions. But they have the explaining power on looking at the problem studied here. There are two points need to be noted:

firstly, we don't include the technology factor into the setup when generalizing the model. This exclusion is purely a simplification and the result will not change with adding technology factor. Thus in an economy with continuous growth, Proposition 1 will not change;

secondly, the sum of the parameters α and γ in the production function $Y_i=K_i^{\alpha}\left(\frac{R_i}{R}\right)^{\gamma}$ can be equal, larger or smaller than 1. This means that the above results are not subject to constant return to scale.

This model excludes labor wage and treats capital revenue as the main income resource. We should be aware of this assumption that it makes the model somewhat different from the real economy. In capitalism economy, inequality is a product of the system as what Marx and many other scholars have argued. Smith (1776) already pointed out that the distribution rule between capitalist and workers are different that the former takes profit and the later earns wage. It is thus can be understood that the result of the model shows a kind of equal consequence without the existence of any

wage-dependent labor. Although the assumption is strict, the equal distribution of rent will be still good for the whole society to be more equally distributed. Stiglitz (1969) argued that if the labor force is heterogeneous, the wealth distribution will be determined by the productivity distribution. Following this argument and the neoclassical rule that the wage is paid by the marginal product, we know that human capital distribution matters here which is further mainly determined by education. The more equal proliferation of education will bring about more equal human capital distribution and further will bring about more equal wage revenue. The model built here enhanced the model of Stiglitz (1969). From both work we can call on the attention to more equal distribution of rent and education to achieve the goal of wealth equalization in the long run.

The more important meaning of the result is to support the neoclassical frame that such frame can lead the world to be more equal which is different from most classical arguments which tend to propose that the income inequality will become increasingly larger (see the review of Bronfenbrenner (1971) Chapter 4)

3.5 Concluding remarks

This chapter studied the income/wealth distribution evolution during economic growth with rent-seeking. We extended Ramsey-Cass-Koopmans model with two representative agents representing "inside institution" and "outside institution" economies. This arrangement is based on the special institutional arrangements of China's economic reality. The model results are consistent with Chinese economic performances that the inequality has been going larger and larger without a Kuznets "inverted U" curve appeared so far.

The first model with two-representative agents showed that the rent plays a determinant role in wealth distribution dynamically. Although the two-agent model is designed based on China's institutional arrangements, the frame is also shedding light on looking at other economies. The second model implies that the "double-track" economy should be transferred to the "single-track" with abolishing the "inside

institution" economy so as to achieve the maximum output. We also generalized the model to study the rent-inequality nexus. The model shows that as time goes to be sufficiently long, rent distribution and wealth distribution are showing a positive and determinant trend that more equal rent distribution will lead to more equal wealth distribution in the neoclassical frame.

The modeling work also reveals that the neoclassical frame which ignores the factor of institution cannot fully and correctly explore the development problems. The factor of rent is adopted to describe the role of institution in economic development. The results achieved in this chapter enhanced the arguments of Stiglitz (1969) and Chatterjee (1994) which expanded the understanding of rent-seeking space in economic development.

Chapter 4

Economic Growth and Transaction Cost

4.1 How to understand economic growth?

4.1.1 Introduction

Economic size is usually measured by GDP. GDP itself is not a difficult concept which simply means the total output and service provided within one year of a country. But how to understand its size change and dynamic growth is far more complicated and difficult than usually thought. People in this world never met constant GDP growth till late 18th century when the “industrial revolution” happened. According to the estimates of Maddison (2006), the annual average compound growth rate of the world per capita GDP is 0.00% from 0 to 1000 A.D., 0.05% from 1000 to 1820 A.D. and 1.21% from 1820 to 1998 A.D. If we focus on the history from 1500 A.D. to 2001 A.D., we have the numbers below:

Table 4.1

Per Capita GDP Growth : World and Major Regions, 1500–2001

(annual average compound growth rates)

	1500-1820	1820-70	1870-1913	1913-50	1950-73	1973-2001
Western Europe	0.14	0.98	1.33	0.76	4.05	1.88
W. Offshoots	0.34	1.41	1.81	1.56	2.45	1.84
Japan	0.09	0.19	1.48	0.88	8.06	2.14
West	0.14	1.06	1.57	1.17	3.72	1.95
Asia (ex Japan)	0.00	–0.10	0.42	–0.10	2.91	3.55
Latin America	0.16	-0.03	1.82	1.43	2.58	0.91

E. Europe & USSR	0.10	0.63	1.18	1.40	3.49	–0.05
Africa	0.00	0.35	0.57	0.92	2.00	0.19
Rest	0.02	0.06	0.82	0.65	2.83	1.75
World	0.05	0.54	1.30	0.88	2.92	1.41

Source: Maddison (2003), pp. 263.

From table 4.1, we can see that between the year 1500 and 1820, world per capita income rose only 0.05 per cent a year. From 1820 to 2001, it averaged 1.23 per cent, nearly 25 times as fast. The history from 1950 to 1973 was obviously much more prosperous than we normally thought: the compound growth rate was close to 3%. The data shows that the economic growth after World War II has been much faster than before. This is the recovery effect which causes quicker convergence rate to the steady state, according to the neoclassical growth theory. There is another story, however, happened in the same period but completely different: even in the official report of Chinese government, around 1960 there were tens of millions of Chinese people starved to death. This is called “three-year natural disaster” in China which actually is caused by the failure of economic institution and policies. From 1952 to 1978, as introduced in Chapter 1, the average GDP growth rate in China was still 4.39%, according to Maddison (2007). It is thus obvious that economic growth is just one aspect of an economy. The distribution matters as well. The relationship between growth and distribution is thus an important question for development economics.

We discussed the effect of growth on distribution in previous chapters. Before we move on to understand the effect of income/wealth inequality on economic growth, we need to explore the economic logics behind growth. For the phenomenon described above, one consistent theory is needed to explain all the levels and changes. In addition to explain why the global economy took off after Industrial Revolution (around 1820s)，we also need to understand why we can have the continuous high growth rate after 1820 and how to explain the regional different growth performances since then.

There are too many economic theories to explain what happened and they usually

criticize or ignore each other. But it does not mean that the efforts made are wasteful. They do give the clues and logics toward the “perfect” answer. I contend that the full-picture behind the growth phenomenon should be viewed from a comprehensive angle. In the next section, I will introduce and comment on several important growth theories as a background for my further analysis on growth-inequality nexus.

4.1.2 A reflection on the theories of economic growth

It is not easy to comment on the main schools of economic growth. The difficulty comes from two ways: one is that different schools often ignore each other school's logic and methodology which makes each story usually unrelated. This phenomenon also happens within the same school. Different beliefs make economists walk on their own ways. Thus the arguments and models of such schools were not thoroughly discussed and corrected. This result makes the overall comprehensive understanding on economic growth difficult. The other problem is that the fundamental questions of analyzing growth were not scientifically solved, like what kind of production function is scientific for dynamic analyzing. Most economists just follow the tradition and use it without checking the ground.

Based on the above two problems, I will in the following section try to offer a brief review with rethinking the mainstream theories and give comments. The numerous empirical studies on these schools will be by the large omitted, for the reason of space limitation plus that the empirical results are even more mixed for varieties of econometric techniques adopted.

Domar's Work

“Harrod-Domar Model” has been treated as the beginning of the modern period of explaining economic growth by dynamic models. Before this model, economists were focusing on static theories and rare mathematical models had been created to explore the problem of growth. Of course it does not mean that economists before Harrod and Domar didn’t offer any explanation on economic growth. Actually the founder of economics Adam Smith already argued the idea of wealth creation by increasing

division of labor with productivity advance. “Harrod-Domar Model” is treated as the beginning of the modern period mainly because its method and argument became the foundation of later main stream research: the neoclassical growth theories.

However, there is something confusing here in the history of economic theory. Although the models developed by Harrod (1939) and Domar (1947) were called one name later, actually their arguments were not totally similar. Thus when in 1956, Prof. Solow criticized many points of the “Harrod-Domar Model”, actually some of the points he criticized were not appeared in Domar’s paper and he used the condition which Domar criticized. I would like to point out the importance of Domar’s model in 1947 since his talented idea has been continuously ignored by most literatures in development economics.

The key importance of Domar’s paper comes from the argument of the first part which is usually ignored (later textbooks often focus on the 3rd part of his paper which is about the dual nature of investment). Domar began his paper in asking the question: is the absence of hoarding (of the demand) also a sufficient condition for full employment? He does not doubt the “necessary condition” but the “sufficient”. The question he put forward was based on Keynes who thought that if “S=I”, then full employment will maintain if started from full employment. Domar’s argument proves that “S=I” is not a sufficient condition to maintain full employment in reality, even from the status of full employment[2]. This means that the condition “S=I” isn’t suitable for dynamic equilibrium analysis. I would like to argue that this point has a key meaning since almost all the theoretical works after his paper ignored this point and continued to use the condition of “S=I” as the capital formulation in equilibrium dynamic growth process. Strangely, I didn’t see any criticism on Domar’s argument on this point in the later history or the second time to rethink the condition's effect in a dynamic process.

Solow (1956) criticized that “Harrod-Domar Model” is based on the “*crucial assumption that... there is one possibility of substituting labor for capital in production*”. But in Domar’s paper, he even used a full paragraph to say that “capital may be substituted for labor”.

Domar repeatedly mentioned in his paper that the balance between demand and supply can be unstable thus neither technology nor saving, "guarantee a rise in income". This argument, unfortunately, is also ignored by most following researches. Domar's logic on the well-known "dual nature of the investment process" can be expressed by:

$$\Delta I \cdot \frac{1}{\alpha} = I \cdot \sigma \qquad (1)$$

The left side is the demand side which takes effects from a multiplier process. The α is the marginal propensity to save and the σ is what he called "the potential social average productivity of investment" which includes the capital depreciations and relative effects. The above equation clearly shows that the investment must expand continuously to ensure the full employment since the equation (1) means: $\frac{\Delta I}{I} = \alpha\sigma$ where the right side is always positive in reality.

Later researchers usually criticize the above result by a so-called "knife-edge" problem which is that if the real investment speed exceeds the required level, the demand side will exceed the supply side. Thus when the companies are making too much investment, they will face the dilemma of supply shortage and this information will further lead to more investment and the circle will go on, which means the economy will never go back to the balance. Although I don't think Domar's investment formula $\Delta I \cdot \frac{1}{\alpha} = I \cdot \sigma$ is absolutely correct, I have to say that the above criticism is quite weak. Actually before they criticize after using many mathematical calculations, we can see this result from the equation structure: the left side is a multiplier effect and its realization speed must be faster than the right side. If criticizing that the demand side expansion by investment should be slower than the supply side, we need further proof which I didn't see. Further, since the price will rise for the shortage, the demand power will be checked.

The main problem of "Harrod-Domar Model", in my view, stays in that the dual effects of investment cannot keep balanced in the long-run. Although in the short run, the demand effect can be equal to the supply effect, the demand effect will disappear and the supply effect will continue in the next period because of the increment of capital. The Domar model explains the growth effect of investment, putting it in a balance

frame. He does not explore further that when the balance is broken, how the economy goes back or how long the economy can survive to grow under such a break.

Exogenous growth theory: Solow Model

"Solow Model" is the key model in modern research of economic growth which is also the start of the exogenous growth theory. Solow model differs Domar model at a crucial point that he takes the condition "I=S" as the capital evolution process. This means he totally ignored what is argued by Domar that the condition "I=S" is not a sufficient condition to sustain the economy of full employment. I don't think Solow was aware of this when he was writing this paper because he did not spend any word on this. Unfortunately, all the later literatures also followed Solow's logic and didn't pay attention to this point. The later economists commented that it is because the new production function Solow applied that overcomes the "knife-edge" problem of "Harrod-Domar Model". The function he applied is the commonly used one: Y=F(K, L). And this function is often revised as Y=F(K, AL) which is called "Harrod neutral" to incorporate the factor of technology into the production function, keeping the existence of steady state growth.

Solow model and its relative later models show that the economic growth rate will converge to a steady state where the ratio of capital to labor or augmented labor (K/L or K/AL) will converge to a constant value (or constant growth rate), wherever the starting point is. The assumption is the concavity of the production function, or the production is constant return to scale. After just applying rather simple mathematical transition to K/L or K/AL (total derivatives), we can see the convergence result. After the convergence happens, only technology progress determines the growth rate in the long run. The original Solow model showed that there are two variables determining the steady state of country: the rate of saving and population growth.

As is well known, the concavity of the production function of the whole economy was seriously criticized by the "Endogenous Growth" economists like Paul Romer (1986, 1990). Romer put forward the idea that the whole economy shows an increasing return to scale and he thought that the main problem of Solow model comes from the

concavity assumption of production function. Let me leave this important debate to later discussion and look at other problems in the frame of Solow.

I used the neoclassical production function in the dissertation but we have to admit that the production function F(K, L) is a myth. Domar did not use the production function. There are many debates and discussions on the simple-looking function in economic history. Even a big event called "Cambridge-Cambridge Controversy" was mainly focusing on the question that whether the production function is scientific and the debate results revealed many problems of production function. The controversy is mainly concerning static economics. Actually Solow was a member in the controversy and he realized the so-called "Wicksell Problem" behind the function and thus used the assumption of "only one commodity" in his model (it is already unrealistic). To repeat the details of this long time controversy is not the task of the dissertation. What I what to point out is that, even under the "one commodity" assumption, the production function is still not a scientific instrument in arguing economic growth process when combined with the condition "I=S".

The crucial reason is as what Domar pointed out that the capital formulation from the condition "I=S" is not sufficient to ensure full employment in the next period. So if using the formula: $\dot{K}(t) = sY(t)$ to describe the capital formulation process (it does not matter whether you put depreciation effect in the equation), then the L in the production function of next period will be smaller, if we keep the population growth rate as zero. After periods, the unemployment rate will be higher and higher and the growth process will be seriously damaged. But this dynamic effect has been totally ignored if we use the production function.

This is an example of contradiction between different economic schools. The neoclassical growth model which is built on Walrasian Equilibrium frame, assumes the full employment in any economic dynamic process. Domar, however, did not build his frame under such strict assumptions but based on the real economy.

Let us recheck the production function Y=F(K, L). K and L are inputs including capital and labor. F(•) is a function and its meaning can be abstract. F(•), in my view, can be

seen as both a kind of institution and mechanism within which the capital and labor are arranged although in neoclassical economics it has no relationship with institution but only a math formula like Cobb-Douglas function. What I want to emphasize is that F(•) should be changing during long-run economic growth. But we never consider the change of F(•) when we do dynamic analysis. In 1980s, some social scientists and political economists began to study the supply side of modern industrial economies. Streeck (1992) summarized one of their main arguments that different institutional conditions may cause different production patterns that may represent functionally alternative.

Ramsey-Cass-Koopmans Model

I already applied the Ramsey-Cass-Koopmans model in Chapter 3 to discuss the evolution of wealth distribution. The main improvement of this model compared to Solow's is the explanation for saving rate which is exogenous in Solow model. The Ramsey-Cass-Koopmans model applies the optimization method to get the consumption and saving paths. Thus the model becomes more general and also more complicated. The technology progress is also the growth impetus which is still exogenous in the model. The steady state of the model is saddle-path stable. There are some technical differences between the Ramsey-Cass-Koopmans model and the Solow model, but both of them cannot explain the growth-determining factor: technology progress. Thus the later development has been focusing on modeling the technology progress which is called the "endogenous growth" models.

Endogenous Models: Romer

The school of endogenous growth has been the major school in understanding economic growth since Romer's paper in 1986. Till today, it is still the main stream in economic growth, although many economists strongly criticize this school from many aspects. The representatives of this school are Romer, Lucas, Helpman, Aghion, etc. In this part I mainly refer to Romer's work.

The paper of Romer (1986) is based on his Ph.D thesis which firstly presented a formal

model showing that the production function exhibits increasing marginal productivity of aggregate knowledge. The elegance of his model is that the production function is continue to be concave to its own inputs and private knowledge but convex to the aggregate knowledge where he introduced the role of increasing return to scale to the model. This knowledge is separated from physical capital. He did not introduce the concept of human capital in that paper but we know that the knowledge can only understood by human. The technology in his model depends on the aggregate knowledge level. So he found a way to make the variable of technology endogenous. Romer not only contributed to the progress of new ideas but also to the model designs. Romer's model assumes that the capital change is a function of the ratio of investment to private knowledge (only mastered by a firm). He gave up the tradition that capital change should be equal to new investment which originated by Ramsey (1928). Romer (1990) deepened his argument in 1986 and made the description of technology more deliberate. He treated technology component A as a non-rival input and H, the human capital as the rival component of knowledge. Further, he assumed that the technology evolves as a function of human capital employed in research. This model finally shows that population size does not have much relationship with growth, but the stock of human capital determines the growth rate. This is maybe the weakest point that often induces criticism that it means that the advanced economy with higher education level should always grow faster than the less advanced economy with much lower human capital. This is not the reality and thus we must find the answer for the mismatch of the model. From the point of my view, the above result roots in the assumption that $\dot{A} = \delta H_A A$. Base on this equation, the stock of the human capital employed in research determines the change of technology. This argument may be reasonable to the most advanced economies but not for all economies since technology innovation can be imported from abroad without R&D and thus the human capital may not play a crucial role in technology updating. This paper, however, is a very important start for later models endogenizing technology progress with a simple way. Romer was thus able to influence two branches in endogenous growth theories with the two papers discussed.

The Reversal of Exogenous School

Although the rising of endogenous growth school looks very promising, the insistence on exogenous models never quit. One important reason why economists doubt the exogenous models are not the ones I proposed above (like the production function or the dynamic full employment condition), but their theoretical predictions were understood as against the reality that the convergence in income variation in different countries never happened. But this belief has also been continuously challenged.

Mankiw, Romer and Weil (1992) added human capital to Solow Model and empirically test it. They changed the production function to

$$Y(t) = K(t)^{\alpha} H(t)^{\beta} (A(t)L(t))^{1-\alpha-\beta}$$

H is the stock of human capital.

Their test result on this simply revised exogenous model is surprisingly good: with $\alpha=\beta=1/3$, the model can account for 80% of international variation in income. The convergence also happens according to the rate the augmented Solow model predicts, holding population growth and capital accumulation constant.

This paper has been widely scrutinized for its reliability of econometrics. In my view, the paper also suffers from the problem of unrealistic but important assumptions, e.g. they assumed that all countries have the same depreciation rates and technology innovation rates.

Parente and Prescott (1994) is another strong support for the exogenous models. This paper analyzed the barrier to technology adoption. They argued that it is the difference in this barrier that leads to huge observed income disparity. They proposed that a country's Total Factor Productivity (measure the technology adoption) is a function of its policies and institutions and thus countries have more costly policies and institutions will have lower Total Factor Productivity.

Further, in the paper *The Failure of Endogenous Growth* published in 2001, Parente summarized the past two main schools as below:

My own assessment is that this line of research (endogenous school) has not proven

useful for understanding the most important question faced by economists today, namely, why isn't the world rich. Exogenous growth theory, in contrast, is.

New Institutional Economics: Douglass North

Analysis based on institutions or the so-called "social infrastructure" has become one important branch in explaining economic growth. Douglass North is one of the representative economists of the "New Institutional Economics". In his many contributions, institution is treated as the main variable for economic growth because he argues that the transaction cost (defined by Coase (1960)) is the key to the performance of economies through the channel of increasing productivity, and institution is the determinant of transaction cost. As what he argued below, institution, which is totally ignored by neoclassical economics, plays the crucial role for economic growth:

Modern economic growth results from the development of institutions that permit an economy to realize the gains from specialization and division of labor associated with the sophisticated technology that has developed in the Western world in the last several centuries.

Source: Douglass C. North (1987). Institutions, Transaction Costs and Economic Growth, *Economic Inquiry,* 25, pp. 419-428.

North argued that the institutional arrangement that lower the transaction cost (per exchange) is the promise for productivity progress. He agrees with the point that specialization and division of labor are the inner motivations for growth of national wealth. This point of emphasizing on the role of labor division was further developed by some other economist, like Yang (1991). North further argued that the development of specialization and division promotes the evolution of institutions because it will be required to secure the results of increased trade (this is not the root reason for institutional evolution which will be introduced below).

North also pointed out the historical lesson shows that the political system often brings

inefficient institution innovation and thus causes stagnation. This point, in my view, should draw more attention from governments and further research is required. Then how to judge an institution is a good institution? North said in a speech (2002) that we still know little about institution. He said our knowledge on human institution is still very limited and no one knows what kind of institution is the best for economic growth. Also, we still know little about institutional change. For example, what is the optimal path for an institution to evolve? Economists used to think that the "Shock Therapy" is good for transition from planned economy to market economy and economic history proved that it is wrong for many countries.

According to North's logic, we can say that one of standards to judge whether an institutional innovation is good or not is to check whether the change will lower down the transaction cost. Further, the discussion of this dissertation also points out the importance of rent for distribution. Thus I would like to argue that another standard is to see whether the institutional innovation will subtract the rent-seeking space and make the rent distribution more equal. In addition, the experience of China since reform and opening up shows that the gradual reform path is workable. This philosophy is called "cross the river with touching the stones" (Xiaoping Deng). The gradual path shows the ability of overcoming the problems happened under the "Shock Therapy" path. This is an important topic deserves more studies.

North pointed out that the most crucial motivation for institutional evolution is the fundamental change in relative prices. This is a very sharp view since normally people always relate institution to political game or revolution, etc. North argued that the change in relative prices will require people and organizations to make the new contracts which will change the norms and then the rules of the society. He also pointed out that the changes of ideas and values are also influenced by the changes in relative prices. The institutional innovation, however, is generated by the consideration of comparing the marginal benefit and marginal cost of inducing such a change. We can see from here that when technology updates and productivity increases, the relative prices will change automatically and thus the institutions will update at the same time. This also means that F(•) changes during economic growth.

North’s argument also built one connection between the new institutional economics and classical economics since the relative price is also a key concept in classical economics. Sraffa (1960) and Schefold (1997) showed that the relative prices are a function of the rate of profit in the classical economic frame. Combined with North’s arguments that the relative prices are the root cause of institutional innovation, this implicitly points out that the distribution takes a crucial effect on institutions. The previous chapters of the dissertation, however, already showed the opposite relationship that the institutions shape the distribution. Thus there seems to be a mutual causal effect between them.

Summary

The above ideas and theories are just part of the story on explaining economic growth (I did not review most empirical works and some other growth theories) and we can see that even these few theories have contradictions and many imperfect sides. We have to admit that the space of revising existing stories and creating new ones on economic growth is still huge.

Of course, we can look at this problem more loosely that we may think that economic growth has not only one root. We may think that each of the factors mentioned above like investment, technology, human capital and institution plays the crucial role and we need them as higher or better as possible in a balanced step. But this attitude is not satisfying for really understanding the growth problem and we need a model that can incorporate the ideas with consistent logics.

From the above theories introduced and commented, we can at least get some clues about the determinants of growth:

First of all, the global history since industrial revolution clearly shows that the adoption of modern technology caused the sudden change of economic history. Modern technology is the fundamental root of the largely increased productivity. Without the modern technology, productivity can be increased by several other ways like update management or create a more efficient institution. But these other ways of updating cannot maintain.

Further, new technology is expensive to be discovered and broadcasted. There are two ways for a country to get new technology: through R&D itself or import from other countries. The adoption of the new technology (make it productive) is different from the invention of the new technology. So from the above argument, we can see that investment is a must for growth. Without continuous investment, R&D or technology-import cannot happen continuously. But since the property of diminishing return of capital accumulation, the progress of technology is the only reason for the continuous investment.

Thirdly, Karl Marx and Joseph Schumpeter emphasized the entrepreneurs' effect on continuous technology innovation. The institution of Capitalism is thus important for technology innovation. Acemoglu and Robinson (2012) called the institutions encouraging economic growth "inclusive economic institutions". The perfect competitive market, as supposed to be the best economic system to realize efficiency, directs the development of Capitalism institutions. The importance of an efficient institution reminds us that the government can give a visible hand on updating both the political and economic institutions. Acemoglu and Robinson (2012) pointed out that the "*there is strong synergy between economic and political institutions*". This point can also be seen from our analysis in Chapter 2.

Finally, the human capital, usually measured by education level, seems to have several crucial effects: firstly, it can be one important factor of influencing institutional innovation since institutions, as what North defined, are "rules, enforcements characteristics of rules, and norms of behavior that structure repeated human interactions". Institutions are made by people. Thus people's human capital, in my view, plays a role here. Secondly, human capital obviously influences the possibility of R&D and adoption speed of new technology. From the above logics, human capital is indeed an important variable since it has influences on both the above factors, although it may be not the most crucial factor affecting institutional innovation or technology adoption of a less developed country.

One thing needs to be noted is that the role of the income/wealth distribution didn't draw much attention in most growth theories. As I argued in Chapter 2, such models

usually implicitly assume that the inequality problem can be solved with continuous growth and it is not the crucial factor in growth frame.

4.2 Transaction cost and the size of an economy

4.2.1 Introduction

After reviewing the main schools' economic thoughts on growth, let's come back to the most basic economic factor: transaction (exchange). In economic literatures, transaction has been discussed by some economists. Smith (1776) emphasized the role of labor division on promoting economic efficiency. He pointed out that the division of labor is not a result of human wisdom but from the propensity in human nature to do transactions which will bring them utility. Commons (1934) put forward the proposition that the transaction should be treated as the basic unit of economic analysis. Mao (2008) argued that wealth is created by fair and free transactions. He also pointed out that when both sides agree upon a deal, each side will get utility from the deal if the deal is agreed upon without any unfair factors like pressure or order. The frame of illustrating how transactions take effect on economic performance, however, is still somehow a black box in economics.

Economics field paid more attention to the cost of transaction as what Williamson (1985) said "*transaction cost kept reappearing*". The concept of transaction cost, since its appearance, has been applied much more frequently on the microeconomics field, e.g. the well-known school of transaction cost economics (TCE) developed by Williamson (1975, 1985, 1995) and some other economists. TCE is mainly concerned about economic organization. Although the new institutional economics is a try to include the neoclassical resources into the institutional analysis, the marginal and equilibrium analysis in TCE are very rare. The main way TCE does is to implement comparative institutional assessment of discrete institutional alternatives. North (1973, 1987, 1989) applied the concept of transaction cost to explain economic growth and pointed out that transaction cost is crucial for the growth performance of an economy and institution is the determinant of transaction cost. His argument is established from

a long-run historical perspective and most of his effort is focusing on the role of institution. Although North's effort updated our understanding on the role of transaction cost, the formal model on how transaction related to economic growth has not been well established. Yang and Borland (1991) built a model between labor division and economic growth, pointing out that the exchange efficiency affects the evolution of division of labor where they created a relationship between transaction cost and exchange efficiency. Their work on the transaction cost, however, contains two problems: the first is that the transaction cost defined in their paper is not exactly the traditional concept developed by Coase and Williamson; the second problem is that under their structure, the transaction cost will increase as the number of trade partners increases. At least, this assumption is too strong and not quite clear either. Besides, the researches on how to apply the concept of transaction cost to analyze the short-run macroeconomic fluctuations or the changes of the size of an economy are even more difficult to be found in literatures. Especially, "positive" transaction cost has always been treated as outside of the neoclassical main stream economics and ignored in most modeling works based on general equilibrium frame.

I would like to argue that one of the crucial reasons for such incompleteness and neglect comes from the abstract and complicated meaning of the transaction cost. This concept was initially set up by Coase (1937, 1960) but gradually developed into similar but diversified understandings. Since Coase himself didn't offer the name of "transaction cost" and his definition on it is somehow open to discussion, later economists further gave many definitions of this basic concept in new institutional economics. Arrow defined the concept as the "cost of running the economic system" (Williamson 1985, pp.18). For Arrow, transaction cost is like the friction in physical system which is the smaller the better. Williamson (1975, 1985, 1995) developed the definition of transaction cost, dividing it into several different kinds of costs, like search and information costs, bargaining costs, policing and enforcement costs, etc. Williamson (1985) gave a detailed definition of transaction cost which include ex ante and ex post types related to contract. Kasper and Streit (1998) defined that "*transaction cost are the costs of employing recourses when people use markets to*

exchange property rights". Their definition is similar to Williamson's. Further, Cheung (1998) argued that transaction cost is actually the "institutional cost" which measures all the costs which do not exist in a Robinson Crusoe economy. Obviously, his definition is much broader for the reason as he argued that it is usually impossible to differentiate one type of transaction cost from another. Under the transaction cost concept of Cheung (1998), the more advanced the economy, the higher transaction cost will become and transaction cost is not the reason behind economic growth. To the opposite, North (1989) treated the transaction cost has an inverse relationship with the prosperity of an economy. I will point out later that the confusion here can be clarified by the concept "unit transaction cost".

Such abstract and diversified definitions make the concept hard to appear in neoclassical works which always assume the non-existence or zero quantity of transaction cost. Besides the definition reason, the new institutional economics, esp. the TCE, assumes that people have bounded rationality which is one background of the existence of transaction cost. For them, this cost is also closely related to human nature. This is also different from the neoclassical assumptions.

I would like argue that the concept of transaction cost and transaction actions have some unique powers in explaining economic performance. The analysis based on them can not only avoid using both the production function and the full employment condition "I=S" in the analysis of economic growth, but also can analyze institutions like what North did and should be more applicable for a transitional economy without the institutional arrangements required in neoclassical equilibrium frame. State and customary institutions, as argued by Hodgson (1988), are required for a system of transactions. Williamson (1985) concluded that the institutions have the main purpose and effect of economizing the transaction cost.

This section is devoted to enhance our understanding on the effects of transactions in economic performance. We will fulfill this goal by modifying the traditional concept of transaction cost. The revised concept of transaction cost in this section is a simpler one which mainly concerns the Williamson "ex ante" type of transaction cost (but not totally the same). The relationship between transaction (exchange) and the size of an

economy both in the short-run and long-run is studied with the new transaction cost. For simplicity, we assume a closed economy for the theoretical approach.

4.2.2 The basic model

4.2.2.1 The transaction cost formula with a newly defined concept

I apply the concept of transaction cost to describe the feeling "easy" or "hard" to achieve an acceptable transaction (or exchange). The lower the cost means the easier to make the transaction done. But how can we describe the feeling "easier"? I assume that any transaction cost is a combination of transaction time, labor efforts and money cost (exclude final transaction price which is the price of the final good or deal) with different weights. The transaction time, transaction labor effort and money cost are the amount spent from the beginning of desiring to do a transaction till the end of the transaction completed.

DEFINITION 1. The "transaction cost" is the cost of achieving an acceptable transaction which is a combination of transaction time, transaction labor effort and the money cost to complete the transaction (exclude the price of the final good or deal) with different weights.

Denoting the definitions of the three factors in achieving an acceptable transaction as below:

Time (T): the transaction time from the beginning of desiring to do the transaction till the end of completion;

Labor (L): the labor effort spent from the beginning of desiring to do a transaction till the end of completion;

Money cost (C): the money spent in achieving the transaction (exclude the final price) from the beginning of desiring to do a transaction till the end of completion.

Here we assume that the desiring is accompanied by immediate action to pursue the transaction.

If we let the T(t) mean the transaction time from the initial time t to t+T(t) with L and

C the functions of the transaction time where L(s) and C(s) means that at the time point s, the labor effort and money spent to achieve the transaction. Based on the above definition, we can express the transaction cost (TC) as:

$$TC(t) = \int_{t}^{T(t)+t} [w_1 \cdot L(s) + w_2 \cdot C(s)]\, ds \qquad (1)$$

where w_1, w_2 denote the weights of labor and money cost contained in a transaction. Similar to the production function, there are many possible forms to describe the relationship between L and C (like L and K in production function), the form presented here is just a simple general case of the possibilities and it won't influence our analysis on the relative properties.

The traditional concept of transaction cost, firstly put forward by Coase (1937, 1960) and later developed by many other economists like Williamson (1975, 1985, 1995), has more contents than the concept I defined here. What I hope to focus is the cost of achieving or realizing an exchange or transaction and thus I intentionally omit some traditional meanings of the transaction cost such as the ongoing monitoring costs and the implicit costs of contract incompleteness. There is no intention in my definition that such costs can be fully avoided by the efforts during the process of achieving the transaction. Later analysis of the section will show that the simplification of the traditional concept of transaction cost by focusing on the process of transaction achievement will enable us to build a channel between transaction cost and economic size as well as some other issues. Further, although the aspect of uncertainty is not included in the definition, it is already included in the frame since rational people will take all the uncertain factors into consideration and the effort for this is included in the factor of labor and money spent and the cost is also reflected by the transaction time length.

About the unit problem, labor and money costs seem to have different units. We can adjust this difference by measuring the labor cost as the money he can earn like the wage revenue if the same amount of labor is used alternatively. Thus the final unit of formula (1) can be treated as the money unit.

4.2.2.2 Properties of the factors in the "new" transaction cost

Intuitively, we can assume several properties of the three factors in TC(t) defined in the above section:

1. The diminishing speed reduction of transaction time with the increase of transaction labor or money cost or the labor and money costs combined. That is, $\frac{\partial T}{\partial L} < 0$ and $\frac{\partial T}{\partial C} < 0$ with $\frac{\partial^2 T}{\partial L^2} > 0$ and $\frac{\partial^2 T}{\partial C^2} > 0$. This property is opposite to the concavity assumption of production function.

2. There are both substitution effect and complementary effect between the labor cost and money cost. That is, $\frac{dL}{dC} < 0$ and $\frac{\partial^2 TC(t)}{\partial C \, \partial L} < 0$ (this is opposite to production function).

3. The properties of TC(t) has similar characteristics to neoclassical production function F(K, L). That is, $\frac{\partial TC(t)}{\partial L} > 0, \frac{\partial^2 TC(t)}{\partial L^2} < 0; \frac{\partial TC(t)}{\partial C} > 0, \frac{\partial^2 TC(t)}{\partial C^2} < 0.$

The traditional production functions do not contain the factor of transformation time. The transaction time, however, is introduced into the TC function.

4.2.2.3 Transaction volume and the "Unit Transaction Cost"

Suppose there are n economic agents and let each agent's wealth endowment (exclude the price of the final good) for pursuing a transaction is w_i. Define the transaction volume as the amount of money transacted and x_i where $i \in N$ to be the volume of transactions of agent i within a period and define TC to be the transaction cost contained in each unit of the transaction which means 1 dollar or 1 unit of other kinds of money used. We call the transaction cost contained in the one unit of money as the "Unit Transaction Cost".

Based on the arguments of Mao (2008), we assume that more transactions will bring higher utility which means that people hope to transact more since they can get higher utility. Let each agent's utility function be $U = x_i^{\alpha}$, where $0 < \alpha < 1$. With the above setup, we have $x_i TC \leq w_i$. Let λ be the Lagrange multiplier, we can construct the Lagrangian formula:

$$L = x_i^{\alpha} - \lambda(x_i TC - w_i)$$

from the first order condition we have

$$x_i = \left(\frac{\lambda TC}{\alpha}\right)^{\frac{1}{\alpha-1}}$$

So there is an inverse relationship between x_i and TC which means that when the unit transaction cost becomes lower, people will transact more.

The concept of the Unit Transaction Cost has been established in this section which clarifies the delicate difference between the unit transaction cost and the total transaction cost. This difference proves very crucial in later arguments. The traditional economics concerning transaction cost didn't differentiate the two concepts. Thus some economists like Cheung have the opinions that the more advanced the economy, the higher transaction cost. And some other economists like North argued in the opposite direction. I would like to argue that implicitly Cheung is pointing to the total transaction cost and North's argument is based on the unit transaction cost, although he argued it as the "per exchange" transaction cost (1987).

4.2.2.4 Unit transaction cost and the size of a closed economy

Following the above notations, let's denote TC_i be the unit transaction cost faced by an economic element "i" within a certain period (strictly speaking, it should be a certain time point since the transaction cost will be different at different time points, but for simplicity I omit the difference here) and using x_i to represent the total volume of the transactions made by the one element in the economy (include people, companies, governments and other organizations) within a period (e.g. for GDP the time measured is one year). Denote $n<\infty$ describing the number of all the economic elements, so the transaction volumes of all economic elements in a closed economy can be described by the vector $X=[x_1, x_2, x_3....x_n]$. Plus, according to the above section, we assume that each x_i is a monotone decreasing function of each related TC_i which means that when the unit transaction cost becomes lower, the element's transaction volume increases during the time. The function is like: $x_i=f_i(TC_i)$.

If we measure the size of an economy by its GDP, with the assumptions above, the

relationship between the unit transaction cost and the size of an economy can be easily seen from the standard calculation method of GDP.

LEMMA 1. Given the assumption that the total volume of the transactions made by one economic element within a certain period is a monotone decreasing function of the related unit transaction cost, within a closed economy, if the unit transaction cost of every element of the economy decreases, the GDP will increase within the same period, meaning the size of the economy expands.

Proof: A closed economy can be divided into three sectors: consumption, investment and government purchase which are made by consumers, companies and governmental organizations. For a closed economy, $GDP=C_d+I_d+G_d$ (d denote domestic).

We have the individual transaction volume $x_i=f_i\ (TC_i)$ where $f_i\ (\bullet)$ is a monotone decreasing function. Suppose that $x_c=x_{c1}+x_{c2}+\ldots+x_{ci}$, $x_I=x_{I1}+x_{I2}+\ldots+x_{Ij}$, $x_G=x_{G1}+x_{G2}+\ldots+x_{Gk}$ where i+j+k=n which represent the total exchange volume in each of the three economic sectors. Thus if every element's unit transaction cost decreases, all of x_c, x_I and x_G will increase. Plus, we can see each sector following a basic function relationship: $C_d=f(x_c)$, $I_d=g(x_I)$, $G_d=h(x_G)$ where f (•), g (•) and h (•) are monotone increasing functions, then GDP will expand since $GDP=C_d+I_d+G_d=f(x_c)+g(x_I)+h(x_G)$ when every element's unit transaction cost is lower.

Q.E.D.

The above result, however, is too strong since we cannot expect that all the economic elements in an economy will face the same direction of the changes of unit transaction costs within the same period. Even in a growing economy, some people and organizations will still face higher unit transaction costs for varieties of reasons. Thus the concept of "the average unit transaction cost" is of more interest. We want to see whether the average value of the unit transaction costs can determine the economic

size as Lemma 1 shows and how it takes effect. The following analysis is to explore the relationship between the average transaction cost and the size of a closed economy.

4.2.2.5 The function of the average unit transaction cost

Transactions are made by persons and organizations. As above, it is supposed that their total number is "n" where $n<\infty$ and if we add all the individual transaction costs within a certain period "TC_i" of all the economic elements together and divide it by the total number "n", we can express the average unit transaction cost of a society (or an economy) by:

$$\frac{TC_1+TC_2+\cdots TC_n}{n}$$

But since the unit transaction costs of people and organizations are very different, we follow the proof of Lemma 1 and divide n into three parts: i consumers, j companies and k governmental organizations. Then we have the average transaction cost of each economic sector as:

$$\frac{TC_1+TC_2+\cdots TC_i}{i} \ , \ \frac{TC_1+TC_2+\cdots TC_j}{j} \ , \ \frac{TC_1+TC_2+\cdots TC_k}{k} \qquad (2)$$

DEFINITION 2. The average unit transaction cost is a combination of the average unit transaction time, the average unit transaction labor effort and the average unit transaction money cost with different weights to complete the unit transaction (exclude the price of the final good or deal).

If we denote the definitions of the three factors in this average unit transaction cost with small letters: Time (τ): the average unit transaction time, Labor (l): the average unit labor and Money cost (c): the average unit money cost.

Similar to equation (1), we can express a proxy for (2) in each sector as:

$$E(t) = \int_t^{\tau(t)+t} [w_1 \cdot l(s) + w_2 \cdot c(s)]ds \qquad (3)$$

w_1, w_2 denote the weights of labor and money cost contained in a unit transaction on average in one sector.

PROPOSITION 1. There is a relationship between the average unit transaction cost of each economic sector and the size of a closed economy measured by its GDP: suppose that every economic element in one of the three sectors of the economy follows a linear monotone decreasing function between transaction volume and unit transaction cost with the same the marginal effect of unit transaction cost on its transaction volume, then when the average unit transaction cost of the sector becomes lower, the part of GDP contributed by this sector will increase and thus when all the average unit transaction costs of the three sectors become lower, the GDP will be higher, that is, the size of the economy expands.

Proof: From the proof of Lemma 1, we know that $C_d=f(x_c)$, $I_d=g(x_I)$, $G_d=h(x_G)$, where f (•), g (•) and h(•) are monotone increasing functions and the individual transaction volume $x_i=f_i(TC_i)$ where f_i (•) is a monotone decreasing function. Also we have $x_c=x_{c1}+x_{c2}+\ldots+x_{ci}$, $x_I=x_{I1}+x_{I2}+\ldots+x_{Ij}$, $x_G=x_{G1}+x_{G2}+\ldots+x_{Gk}$ where i+j+k=n which represent the total transaction volume of each of the three economic sectors with the average transaction cost of one sector represented by

$$\frac{TC_1+TC_2+\cdots TC_i}{i},\ \frac{TC_1+TC_2+\cdots TC_j}{j},\ \frac{TC_1+TC_2+\cdots TC_k}{k} \qquad (2)$$

The question thus becomes: if the average transaction cost of one sector becomes lower, will the sum of x_i in the sector higher? Since we know the weight of different TC in the sector is different, we need to see how the average amount takes effect.

We now assume the function

$$x_i=f_i(TC_i)=aTC_i+b_i \qquad (4)$$

which is a monotone decreasing function. The assumption of the above function is based on two considerations: the first is that the linear form is a simple and general case; the second reason is that $a=\frac{dx_i}{dTC_i}$ which is a marginal value and thus we can assume the change of unit transaction cost within one sector has the same effect on the change of transaction volume. With such setup, when the unit transaction cost changes the same amount, the difference of the transaction volume x_i is determined by the intersection value b_i but not "a".

Now the question can be formulated that if

$\sum_{i=1}^{n} \Delta TC_i \geq 0$ (5), will we have the result that

$\sum_{i=1}^{n} [f_i(TC_i + \Delta TC_i) - f_i(TC_i)] \leq 0$ (6) ?

If the function $f_i(TC_i)$ is linear, then equation (6) will become

$\sum_{i=1}^{n} a\,\Delta TC_i$ (7)

Since the function $f_i(TC_i)$ is a monotone decreasing function, we can assume that

$a \leq 0$ (8)

From (5), (7) and (8), we can get the result of (6). However, if the function $f_i(TC_i)$ is in other forms, like it is not linear or with "a" not fixed, we cannot always get the result of (6) and further additional conditions are needed.

Thus if the average unit transaction cost of the one sector becomes lower, the sum of x_i in the sector will be higher. For the same logic, when the average unit transaction costs of all the three sectors become lower, x_c, x_I and x_G will be higher. Thus GDP will grow since $GDP=C_d+I_d+G_d=f(x_c)+g(x_I)+h(x_G)$ where f (•), g (•) and h(•) are monotone increasing functions.

Q.E.D.

The effect of the average unit transaction cost can also be seen as the effect of the aggregate unit transaction cost. Interestingly, function (4) has the same structure of the well-known "Gorman form" in Microeconomics whose function is also similar. So here is another case to show the bridge from microeconomic foundations to macroeconomic analysis.

The above proof shows that reducing every economic element's unit transaction cost is different from reducing the average unit transaction cost within one economic sector. The latter takes effect conditionally. Although we proved that the average transaction cost takes effect within each sector, if we consider that the function like equation (4) applies to all economic elements of a society, we can also generate the concept of "the average unit transaction cost of a society (or an economy)" and the relative results will be similar to the analysis based on the sector analysis. But this result will be much weaker than the analysis based on economic sectors since it is

more reasonable to assume "a" in the function $x_i=f_i\ (TC_i)\ =aTC_i+b_i$ quite different among people, companies and governments.

We can also see that the concept of the average unit transaction cost is useful to look at the short-run macroeconomic fluctuations. It is clear that many reasons or factors are continuously changing the average unit transaction cost of each economic sector, sometimes positive and sometimes negative, and thus cause the economic fluctuations. We will present another way of looking at economic cycles, esp. crisis, based on factor movement speeds later in the chapter.

4.2.3 Implications for economic growth

4.2.3.1 The Model

We construct a simple two-period overlapping-generations (OLG) model to illustrate the problem dynamically. Assume that the economy has a representative individual living for two periods, in the first period she is young and in the second period she is old. Let her preferences be represented by a logarithmic utility function:

$$U_t = \gamma lnx_t + (1-\gamma) lnx_{t+1}$$

x_t is the transaction volume in time t. The parameter γ captures the importance of instantaneous utility in the first relative to instantaneous utility in the second.

Suppose that the individual have no inheritance in the first period and only make revenues in this period when she is young, meaning that she can only make transactions by the revenue. The revenue should cover the whole transaction volume as well as the endowment for achieving the transaction. Then during the second period when she is old, the income left from the first period should cover the total transaction cost and the transaction volume for the 2nd period. No heritage will be left.

Let TC be the average unit transaction cost of a society and Y_t be the income in period t, if we suppose that all the revenue will be spent during her life, then there will be the intertemporal budget constraint:

$$TC_{t+1}x_{t+1} + x_{t+1} + x_t + TC_t x_t = Y_t$$

Now suppose there is a social planner who has the objective to maximize individual's utility function subject to the intertemporal budget constraint. Let μ be the Lagrange

multiplier, the Lagrangian for the this problem reads

$$L = \gamma ln x_t + (1-\gamma) ln x_{t+1} + \mu[Y_t - (TC_{t+1}x_{t+1} + x_{t+1} + x_t + TC_t x_t)]$$

The first-order conditions are

$$L_{x_t} = \gamma (x_t)^{-1} - (1 + TC_t)\mu = 0$$

$$L_{x_{t+1}} = (1-\gamma)(x_{t+1})^{-1} - \mu TC_{t+1} - \mu = 0$$

Resolving the two equations and rearrange, we can get

$$\frac{x_{t+1}}{x_t} = \frac{1-\gamma}{\gamma}\frac{1+TC_t}{1+TC_{t+1}}$$

If we make $\frac{x_{t+1}}{x_t} = g$ to be a proxy of economic growth rate and suppose that the growth rate of unit transaction cost of society is $\frac{TC_{t+1}}{TC_t} = h$, then there is a clear relationship here:

PROPOSITION 2.

The higher the growth rate of average unit transaction cost of a society h, the lower the economic growth rate g, vice versa.

This is a progress of Lemma 1 and Proposition 1. Lemma 1 and Proposition 1 didn't include the factor of dynamic growth. Here the relationship between the growth rate of an economy and the growth rate of the average unit transaction cost of a society has been set up. This confirms and deepens the arguments of the new institutional economics on growth, like North (1973, 1987, 1989).

4.2.3.2 The Long-run perspective

The sector-based analysis in above sections can generate some aggregate long-run perspectives. If we view the phenomenon of economic growth as the long-run expands of the size of an economy, the economic history after industrial revolution shows a trend of reducing the average unit transaction cost for all economic sectors compared to the previous time which means that transactions become easier and easier for all economic sectors. For example, when a person who lived in 500 years

ago, if he hopes to travel 1000 kilometers to buy or sell something or to make an investment, he may have to spend half-year which means that he also spent much labor efforts and money compared to his income. But today, with the airplane, people can travel around within one day, spending little labor and the money is not so huge compared to his earning per month. This is the effect of modern technology on transactions. Another fact is the market development, for the long time before industrial revolution, people just have very few times to go to the market in a year since most of the articles they needed were supplied by themselves and not much surplus existed, thus the exchanges with other people were not so possible and frequent. The relative institutions were also formed according to the production level. For example, in ancient China, the governments set up some marketing places with limited time entry and many limitation policies. For example, at the Tang Dynasty China (AD 618-907), governments set up markets in authorized areas and limited the entry time. Today we even have the internet market which is beyond any ancient people's imagination.

Further, the progress of human capital, technology, institution or any other factors which have greatly influenced economic growth also means that they are causing the average unit transaction costs of all sectors to decline. The argument here does not intend to clarify the role of technology, human capital, institution or other factors in economic growth, but to show logic that all the factors influencing growth are related to the average unit transaction cost of each economic sector.

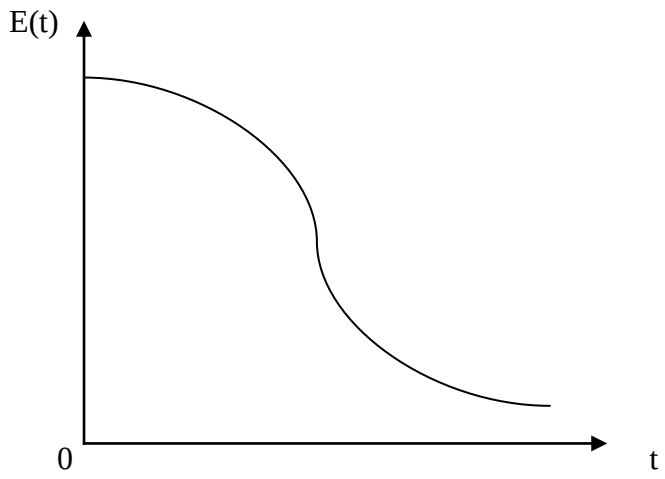
We should pay attention to the changing speed of the average unit transaction cost. Before the industrial revolution when the world was suffering the long "Malthusian stagnation", there was nearly zero growth of global economy and the changing speed of the average unit transaction cost was obviously very slow. After 1820 AD, technology innovations have been keeping improving the exchange instruments which greatly promote the changing speed.

Based on the above observations of economic history, a judgment on the changing trend of the average unit transaction cost of each sector can be outlined. Firstly, we can suppose that the average unit transaction cost of each economic sector has been

reducing from the long-run perspective. Applying E(t) of equation (3) as the average unit transaction cost of one sector, we have $E'(t) < 0$ ($E'(t) = \frac{dE(t)}{dt}$). Further, based on the facts of economic history, it seems reasonable to assume that, from the beginning of human history till around 1820 AD when the industrial revolution happened $E''(t)<0$ and $E''(t)>0$ from the industrial revolution till now. However, we cannot expect that there will be one day the value of E(t) = 0 because it is not possible, so I suppose that there will be a time the changing speed of the average unit transaction cost will slow down and the value of $E''(t)$ will be negative again.

Based on the above arguments, the E(t)-t curve as shown below can describe the trend of the average unit transaction cost for human economic history:

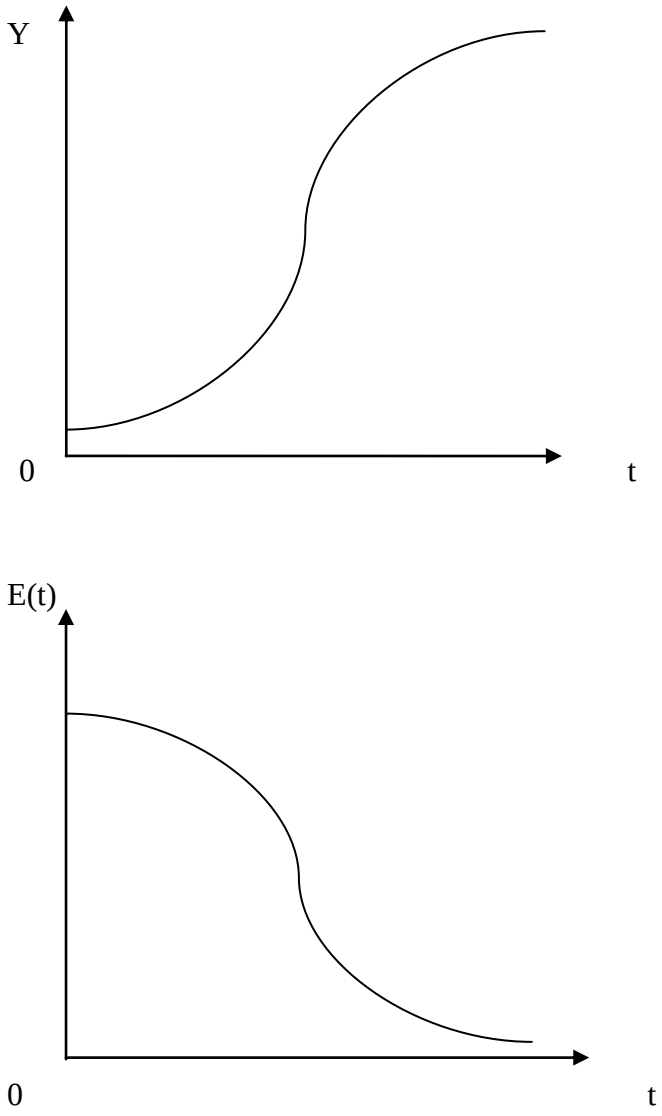
Figure 4.1



As the figure described above, I suppose that the E(t) cannot reach 0 but will go forward to it. Although I don't think there is a limit of science and technology progress, the E(t) cannot be zero except that we don't exist. The curve reflects the declining trend of the average unit transaction cost over history, that is, it slowly declines at the first stage and then it enters a fast-declining period which started from around the 1820's AD and finally the speed will slow down. Further, when the E(t) stops declining sharply, the Y(t) will also enter an area of growing slowly and even

stop growing. Now we are staying at the second period, but it seems to be a very long history. The relationship among Y(t), t and E(t) can be described as:

Figure 4.2

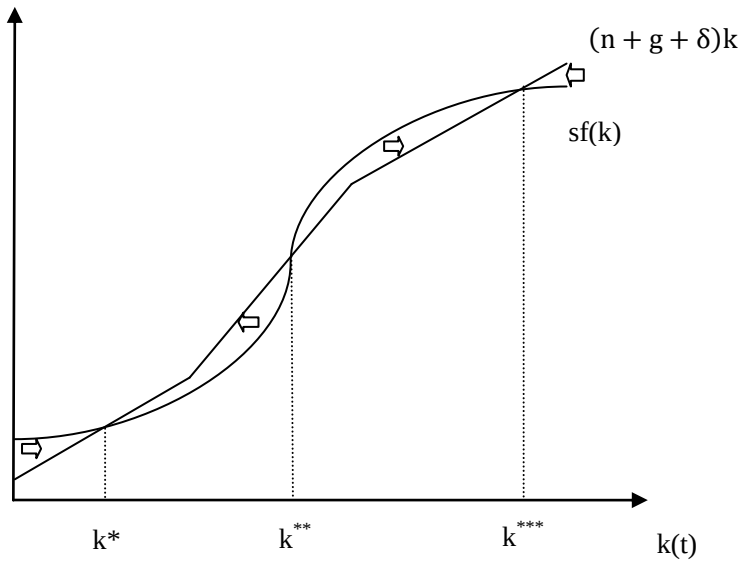


The result of "Figure 4.2" has an interesting similar shape as the production function usually applied by economists. This result helps justify the usage of production function in analyzing economic growth from the long run.

If we follow the traditional production function Y=F(K, AL), based on the shape of Y=F(•) in "Figure 4.2" above, the shape of sf(•) (f(•)=F(K/AL, 1)) will be similar as described in "Figure 4.3". If we follow the logic presented above, the economic

history can be divided into three big time zones. Solow model, as is well known, is the basic foundation for modern growth theories. Following the standard notation of Solow model, we can estimate that the relative values of $n + g + \delta$ (which is the sum of population growth rate, technology growth rate and depreciation rate of capital) in each period are also different. Before the industrial revolution, the three factors are quite small and then it becomes much larger after around 1820 AD. When the economy reaches the third time zone where the changing speed of the average unit transaction cost becomes slow again, it is reasonable to assume that the values of $n + g + \delta$ will be smaller again. Combining the growth pattern based on Solow model and the theory presented here, we have "Figure 4.3":

Figure 4.3



Actually "Figure 4.3" is quite similar to the "Figure II" of Solow's 1956 paper. Solow also pointed out that k* and k*** are two steady states of capital accumulation and k** is locally unstable. However, Solow and the later economists usually apply the production function to test convergence within a relatively shorter period. This is not shown in "Figure 4.3" where the convergence k*** only happens in the distant future. It also shows that in the ancient economic history we also had a steady state point (the

"poverty trap") but it is transformed by the industrial revolution. Thus concerning the current period, I contend to agree with the arguments of increasing returns of production of output by Romer (1986) based on the argument in this section.

4.2.4 Implication for the issues of economic crisis

4.2.4.1 A brief introduction

Economic crisis is always accompanied with capitalism. It is reported that hundreds of economic crisis have happened since industrial revolution. The 2008 financial crisis was followed by credit crunch and has caused huge damage to the economy. The negative effect continues till now with the global economy frustrating with permissive expectations. There are many economic discussions on this crisis blaming on the neoclassical economics and thus Keynes' theory again attracts much more attention. The policies based on Keynesianism or the New Keynesianism, however, performed not so well that the global economy is still far from a prosperous economic upturn.

One of the phenomenon happened during the crisis is that capital firstly flows much faster (escaping) and then changed to be very slow with bad expectations (waiting) while people are harder to find jobs and the commodity market suffers from fewer transactions. In this section I would like to argue that the above phenomenon of factors' flowing has a crucial effect on shaping the crisis and the relative policy targets should also be based on the specific root. This viewpoint is originated from the theory of unit transaction cost presented above.

4.2.4.2 Transaction speed and economic crisis

Supposing that during the normal time without economic crisis, the economy is staying at equilibrium. On such equilibrium points, the demand side is equal to the supply side. When economic crisis happens, the former equilibrium is broken and then the economy comes back to a new equilibrium after some time.

Let the demand side be Y_t, and the supply side be $F(K_t, L_t)$. $Y_t = F(K_t, L_t) = AK_t^{\alpha}L_t^{1-\alpha}$ can also be seen as the equilibrium state of an economy when the total

demand equals to the total supply. This is one of the constant return to scale Cobb-Douglas production functions.

Based on the definition of transaction cost defined in this chapter, we know that Y, K and L which representing commodities, capital and labor have different unit transaction costs which lead to different transaction speed. We define $\frac{dY}{dt}, \frac{dK}{dt}$ and $\frac{dL}{dt}$ as the transaction speeds of the three economic elements in an economy.

From $Y_t = F(K_t, L_t) = AK_t^{\alpha} L_t^{1-\alpha}$, we have

$\frac{dY_t}{dt} = \alpha A L_t^{1-\alpha} K_t^{\alpha-1} \frac{dK_t}{dt} + (1-\alpha) A K_t^{\alpha} L_t^{-\alpha} \frac{dL_t}{dt}$ with $0 < \alpha < 1$, A>0, K>0, L>0.

The above equation shows the equilibrium transaction speeds of the commodities, capital and labor. The transaction speed of Y is a weighted sum of the transaction speeds of K and L.

The potential problem of the above relationship stability comes from the speed "accelerations" of $\frac{dY_t}{dt}$, $\frac{dK_t}{dt}$ and $\frac{dL_t}{dt}$ which are measured by their second derivatives: $\frac{d^2K_t}{dt^2}$, $\frac{d^2Y_t}{dt^2}$ and $\frac{d^2L_t}{dt^2}$.

Let's simply divide the economic crisis into two kinds: financially produced and physically produced. In the 1st scenario, economic crisis begins with fast capital flowing and in the 2nd scenario it begins with fast commodity production change. The 1st kind of crisis we call it financial crisis when financial organizations first step into serious crisis and the 2nd kind we call it "over production" crisis when the enterprises go bankruptcy first.

Thus we can look at the economic crisis from two angles:

During financial crisis with bubble burst, we will have the relationship: $\frac{d^2K_t}{dt^2}$ increases suddenly, leading to $\frac{d^2K_t}{dt^2} > \frac{d^2Y_t}{dt^2}$ which further leads to $\frac{dY_t}{dt} < \alpha A L_t^{1-\alpha} K_t^{\alpha-1} \frac{dK_t}{dt} + (1-\alpha) A K_t^{\alpha} L_t^{-\alpha} \frac{dL_t}{dt}$ and thus the former equilibrium will be broken.

During the crisis made by over production capacity, we will have:

$\frac{d^2Y_t}{dt^2}$ decreases suddenly, leading to $\frac{d^2K_t}{dt^2} > \frac{d^2Y_t}{dt^2}$ which further leads to $\frac{dY_t}{dt} <$

$\alpha AL_t^{1-\alpha}K_t^{\alpha-1}\frac{dK_t}{dt} + (1-\alpha)AK_t^{\alpha}L_t^{-\alpha}\frac{dL_t}{dt}$ and thus the former equilibrium will be broken.

But the above two processes are based on completely different causalities.

Also, since the capital transaction speed and the labor market transaction speed will slow down after the crisis begins, the inequality direction will change into the opposite one: $\frac{dY_t}{dt} > \alpha AL_t^{1-\alpha}K_t^{\alpha-1}\frac{dK_t}{dt} + (1-\alpha)AK_t^{\alpha}L_t^{-\alpha}\frac{dL_t}{dt}$.

Let's look at the current stimulating policies during crisis: obviously there should be a difference between the early happening period and the later crisis period. At the early period when the economy is beginning to break the equilibrium, the policy should be designed based on the two scenarios: the first scenario is the financial crisis when we should try to "cool" the fast investment transaction speed; the second scenario is the over-production crisis when we should either try to raise the commodity transaction speed or to lower down the capital transaction speed. For the first scenario, the method of raising commodity transaction speed is the wrong step because the problem comes from unhealthy investment "bubbles". There are some ways of promoting commodity transaction speed during the second scenario like more opening the market or reduce the consumption taxes.

The government could also do these before the crisis in order to avoid if it can observe the "bubble" or over-production phenomenon in advance. The in-advance treatment will be more effective since some of the policy options will lose their abilities after the crisis happened. For example, the transmission mechanism of monetary policy during crisis probably has broken down and the fiscal policy will probably face the worse fiscal condition of the governments. This argument supports governmental intervening in the economy when it observes that the equilibrium is going to be broken or the economy is stepping into crisis.

We will find that adopting Keynes' policies at the early happening period of crisis will be harmful since it is exactly the opposite effect will be generated, that is, the capital flowing speed will be enhanced by such policies like reducing interest rate when what we need should the opposite: restrain the capital flowing speed with higher interest

rate.

When the crisis comes to the later stage when the commodity transaction speed becomes faster, Keynesianism policies can make a positive effect on economy by promoting the capital transaction speed. I would like to oppose to reduce the commodity transaction speed during the stage since it is not where the real problem lies.

How to judge the first period and the second period during economic crisis is a question we need to answer. In my view this is not so difficult to observe for the government. This analysis also reminds us the danger of “hot money” and the importance of monitoring the capital account.

The change of “L” can also lead to crisis but such scenarios are rare to be found. I would like to argue that the more opening and freedom for laborers will be very beneficial for the long-term growth since it will naturally promote commodity transactions.

4.2.5 Concluding remarks

In this section of the chapter, we extended the analysis of the effect of transaction on economic size from the perspective of transaction cost. The concept of unit transaction cost has been established. We also discussed how the average unit transaction cost takes effect. Then we explored the implications of the theory presented for the long-run perspective and the issues of economic crisis.

There are some future works worth trying, e.g. expanding the model into an open economy and testing the effects of traditionally accepted growth-enhancing factors like technology progress and human capital accumulation on the change of average unit transaction cost of an economy.

Chapter 5

The Effect of Income/Wealth Inequality on Economic Growth: A Rent Perspective

In this chapter we discuss another side of growth-inequality nexus: the effect of income/wealth inequality on economic growth based on the analyses established in previous chapters. Different from the problem of inequality evolution during growth, this problem concerns the reversed effect exerted by distribution. So far, neither of the two topics has reached a common sense in economics. We already discussed the first problem in Chapter 2 and 3. The second problem is also a very important economic problem since it determines or influences the political attitude toward inequality. However, the studies on this problem are even more ambiguous in economic literatures. The relative results are much diversified. We will discuss these studies through a comprehensive literature review. Then we will present a descriptive as well as theoretical analysis on this problem based on China's economic performance after reform and opening up. Our aim is to present a consistent and clear logic for the difficult problem. The discussions in this chapter are closely related to previous analysis in the dissertation.

5.1 Literature review

Unlike the pure growth theories, the studies of how inequality affect economic growth lack a well-recognized basic paradigm like Solow model in the growth field. In this section we review the diversified researches with contradictory results on the effect of income/wealth inequality on economic growth from the perspective of the general background which doesn't treat China as the only focus.

Theoretical Studies

Smith (1776) already put forward some arguments on this problem. He said high earning of labor is an advantage for a society since the improvement of the greater part of an economy cannot be seen as a bad thing. And no society can be prosperous with the greater part of an economy lives in poor. Thus his view is to support the negative effect of inequality on growth and he called for a more equally distributed society. The world when Smith lived hadn't experienced any significant growth and thus Smith didn't mention the concept of growth but the accumulation of national wealth. This, however, can be seen as a similar concept as growth. The argument of Smith seems to be a long-run viewpoint on this problem.

Marx (1885) argued that the increasing inequality is the main characteristics of capitalism which will lead to crisis. Although he didn't directly present a relationship between distribution and growth, the logic behind his ideology is that the effect of income inequality is detrimental for growth in the long run.

Although most classical economists like Smith and Marx were claiming the negative effect of high inequality, in 20th century, earlier theories tended to support the positive effect of higher inequality on growth. Kaldor (1955) argued that the propensity to save of the rich (capitalists who earn profit) is higher than that of the poor (workers who earn wage). So the higher degree of inequality can increase the aggregate social savings which tend to increase the total investment. More investment leads to higher growth rate. His analysis is based on functional distribution analysis. The basic formula he used is: $I/Y=(s_p-s_w)P/Y+s_w$ where *I* means investment, *Y* can be seen as the GDP, *s* means the marginal propensity to save, *P* and the footnote *p* mean profit and *w* means wage. We can see from the formula that the larger the profit share, the higher the investment rate.

Another major argument supporting this positive direction is the idea focusing on the role of incentive. Mirrlees (1971) formalized the idea and argued that the incentives for the laborers will be diminished by the equity and thus less effort will be paid. Siebert (1998) contended that the larger income inequality will encourage entrepreneurship, risk taking and innovations. Okun (1975) also argued that there is a trade-off between equity and efficiency. Since efficiency is often viewed as economic growth, this view

also confirms the positive effect of inequality on growth. However, there are two ways of understanding this trade-off. One is that if the government wants to make the national income more equally distributed, the growth speed will be limited. The other is that the higher inequality will be good for the economic growth. The above two angles are similar but different. The first one is a policy analysis which is focusing on the opportunity cost of achieving equity and the second one is not to discuss the policy's effect, although it contains the factor of policy reactions. This delicate difference should be aware of.

Most of the above arguments are inclining to argue a dominant effect made by higher inequality on growth which means that with higher inequality, the growth rate will be higher too. They also didn't differentiate that whether the effect is instantaneous or needs periods to realize.

After years' ignoring the problem in 1980s, in 1990s the problem of determining economic growth is highly welcome again, including the effect of income/wealth distribution. With applying more econometrics and theoretical modeling, the discussions became more diversified in details. For the theoretical discussions, there are still some arguments supporting the positive effect. Galor and Tsiddon (1997a, b) developed two theories to support the stand of positive effect. The first theory is based on that, for a developing economy, a high degree of inequality is a necessity for the growth of human capital which is crucial for the "take off" of the economy. The second paper argued that in early stages of development the inequality is good for growth and long-run equality as well.

Interestingly, at this period more arguments were inclined to support the negative effect of income/wealth inequality again. However, the arguments concerning whether the effect is dominant became mixed. There are three broad categories for this side: credit-market imperfections, political economy and social instability.

Banerjee and Newman (1993) discussed the effect of initial wealth distribution on subsequent growth. In their dynamic frame, the initial wealth distribution decides the occupational choice between workers and entrepreneurs. This is the effect of incentives under the background of imperfect capital market. Since the structure of occupational

choices also decides consumers' behavior, a new wealth distribution will be followed. Results of the dynamic model are not unique. They argued that economies with initial high poor/wealthy ratio are likely to converge to stagnation in the low-wage situation with a possibility of overcoming stagnation particularly with a large middle income class.

Galor and Zeira (1993) built a model to argue that under the conditions of imperfect capital market and indivisibilities in investment in human capital, the initial wealth inequality affects the economic growth both in the short run and in the long run. The logic behind this is that if the imperfect capital market means that the borrowing is costly, only who owns enough initial wealth can invest in human capital. Thus the poor people may have to work as unskilled. As a result, the distribution of initial wealth will influence the subsequent growth. One of the model's results is that an economy with a large initial amount of wealth but with high inequality will end up as a poor economy in the long run.

Aghion and Bolton (1997) built a model with capital market imperfection and moral hazard, proving that the higher inequality, the lower the growth rate. They argued that the inequality reduced the investment opportunities of the poor people. Also the inequality will increase the level of moral hazard of borrowers in capital market. As a result, they argued that redistribution will enhance the economic efficiency.

Alesina and Rodrik (1994) created a political economy model of endogenous growth, applying the median-voter theorem. They found that the greater the inequality of wealth and income, the higher the rate of taxation decided by the voting and thus the lower the growth. The link is the redistributive policies. So the causality in their frame was also between the initial distribution and the subsequent growth. They also made empirical test for the theory. For the difficulty of testing the redistributive policies, they directly tested the effect the initial distribution on the subsequent growth, using data from 1960-1985 for most OECD countries and some developing countries. The results are supportive of their argument that the higher inequality is bad for the subsequent growth.

Persson and Tabellini (1994) formulated a general-equilibrium model with overlapping

generations to argue that the inequality is harmful for growth through the channel of political redistributive decisions. They also empirically tested the argument with a panel data of 9 countries and the result is in accordance with their expectations in the democracies. Their test also shows that equality is good for investment only in the democracies which is contrary to traditional argument that inequality is good for investment.

Besides, there are some studies arguing that social instability may be caused by inequality and then leads to failure of growth. Alesina and Perotti (1996) argued that higher inequality will bring higher crime rate and social instability which will check the normal saving and investment activities and thus be harmful for economic growth.

Some other studies emphasized the role the macroeconomic volatility which is positively related to inequality and negatively related to economic growth. The argument behind this idea is that the volatility will weaken the incentive of investment.

Galor and Moav (2004) developed a new frame in which they modeled economic development as a two-period process: for the early period physical capital accumulation is the growth engine and for the later period human capital accumulation becomes the growth engine. In the first period the inequality exerts a positive effect on growth through the channel of higher saving propensity which is required for growth. During the second period the economic growth is steamed by human capital when the inequality harms the growth under the background of imperfect capital market. Different from the traditional arguments of imperfect capital market, they argued that the deciding factor should be the comparison of marginal return to human capital and physical capital in different development stages.

There are also some literatures focusing on the effect of redistribution plan. Aghion, Caroli and Garcia-Penalosa (1999) made a review on this and pointed out that the traditional view that redistribution policy is harmful for growth is strongly challenged and doubtful both from the recent empirical and theoretical viewpoints. They asserted that redistributing wealth from the richest to the poorest always has an overall positive effect on growth. The mechanism of such effect relies on imperfect capital market.

Empirical Studies

Now let's look at what the empirical studies say. Empirical studies on the effect of inequality on economic growth have started to draw attentions since 1990s with more applications of econometrics. Similar to the theoretical arguments, the existing results are also very mixed and confusing with different econometric techniques applied, different control variables adopted and different data chosen. Generally speaking, the results from cross-section studies often show a negative effect whereas conclusions from panel data studies and time series studies are quite mixed and unclear.

Perotti (1994) traced the effect of inequality on growth to the effect of inequality on investment. He argued that since growth is lead by investment in physical or human capital, it is enough to test the effect of inequality on investment as a proxy. Further he used the income distribution as a proxy for wealth distribution. He doesn't think this is a serious problem. There were three main theories to explain the channel through which the inequality affects growth: imperfections in capital market, voting on fiscal policy and political stability. All these three ideas point to the logic that the inequality is harmful for continuous growth. Perotti induced some results from each of the three theories and tested them individually using the panel data applied in the paper of Alesina and Perroti (1993, 1996). His results doubted the validity of the second theory that endogenous fiscal policies are bad for investment but supported the other two ideas. One thing should be noted is that he used the initial income distribution in 1960 (his data is from 1960 to 1985) as the variable for testing the relationship between inequality and growth. This tradition appears in many later studies which believe that it is the initial inequality takes effect but not the contemporaneous change of inequality.

Birdsall, Ross and Sabot (1995) discussed the issues concerning income inequality and economic growth in East Asian economies which performed quite ideal in both low inequality and high growth rate. They argued that education is crucial for growth and it also reduces income inequality. Thus some policies such as enhancing education quality will not only promote growth but also reduce the inequality. Thus there is a virtuous circle that lower inequality will leads to more investment in education which further promote higher equality. They argued that this is main reason to explain the

East Asian growth miracle.

Deininger and Squire (1998) created a group of panel data of very high quality which has been continuously applied so far. The data they constructed covers 108 countries from 1960s to 1990s. They made a thorough research on the relative issues concerning inequality and growth. Their empirical results confirmed a negative but weak influence from initial income inequality on long-term growth for both developed and developing countries. They found that the effect of initial asset distribution matters more. They also made tests on exploring the channels for the effect on growth from initial inequality, finding support for the explanation based on imperfect financial market but not for the redistributive median-voter-based theory. They also found that it was the education level that most affects investment but failed to establish any independent links between the initial distribution and investment. They argued that schooling is the main channel through which the initial distribution takes effect and the further test showed that there is a negative relationship between schooling and initial wealth distribution. Thus in their frame, the higher initial wealth inequality leads to worse human capital which is bad for economic growth.

Barro (2000) applied a group of panel data to generate the result that higher inequality tends to retard growth in poor countries (with per capita GDP below $2070 in 1985 U.S. Dollars) and promote growth in richer economies. The overall sample effect of inequality on growth, however, is not significant and implies that the partial relationship between each other is close to zero. The data he used also covers roughly 100 countries over 30 years since 1960s. Besides the data difference between Barro (2000) and Deininger and Squire (1998), Barro did not use the initial income or wealth distribution index but directly used the contemporaneous Gini index. His research also shows that the fertility rate plays a role in the relationship between inequality and growth. Concerning the affecting mechanism, Barro's result shows some difference from both Deininger and Squire's and Perroti's. Barro didn't find any solid relations between inequality and investment ratio or human capital. He did not give further explanation on the affecting channel.

Forbes (2000) challenged the proofs that income inequality is bad for economic growth.

She argued that most studies generating negative relations are not robust and containing two potential econometric problems: measurement error in inequality and omitted-variable bias. In order to correct these problems, she applied the data of Deininger and Squire (1998) combined with some other source to retest the relationship, using fixed effects, random effects, Chamberlain'sπ -Matrix procedure and Arellano and Bond's GMM method. Interestingly, she also tested the initial inequality's effect but just one period before growth. Her results show that the income inequality in one period before has a positive effect on the subsequent economic growth only in the short (5 years) and medium terms (less than 10 years). The relationship is not significant for the long run.

Partridge (2005) argued that it is not surprising that the empirical results are mixed concerning the effect of inequality on growth if there are different short-run and long-run responses. He tested the relationship using U.S. state data from 1960 to 2000. After trying varieties of estimation techniques for panel and cross sectional data, he concluded that the inequality is positively related to the long-run growth and the short-run income inequality response is less clear. His result again confirms a positive relationship but stands in an opposite position to Forbes (2000) who argued that the effect only takes place in the short run.

More recent researches transited the panel data researches to time series studies. Gobbin and Rayp (2008) proposed to investigate the inequality-growth relationship in a cointegrated VAR-setting. They resorted to the Johansen cointegration method to avoid the traditional problems of heterogeneity, omitted variables and endogeneity in empirical studies. The data samples are for the countries: Belgium, the US and Finland. The test results are different for each of the three countries and thus they concluded that panel research including many countries is not a good way to study the inequality-growth relationship. Only the US data seems more in accordance with the result of Forbes (2000).

Andrade, Duarte and Simoes (2011) analyzed the effect of inequality on growth in Portugal from 1985 to 2007 by applying VAR and SVAR models. Their results suggest that the effect is negative. But they also showed that the effect of higher inequality is

good for the accumulation of human capital in Portugal during the period and they contended that the negative effect came from the fiscal redistribution.

Discussions

The problem of the effect of income/wealth inequality on economic growth has many confusing angles. First of all, there is a difference between the effect of initial distribution effect on the subsequent or long-run growth performance and the contemporaneous effect of inequality on growth performance. Secondly, the effect of inequality may be different for the short-run and long-run or for different economies. Further, there is a delicate difference between the views of arguing that the income inequality is bad or good for the growth. Since growth rate can be affected by many factors with some factors dominating the effect, one of them means that the effect is dominating, that is, the higher income/wealth inequality will lead to lower or higher growth rate. The other point means that the effect of high inequality is bad only in the sense that it will partly influence the growth performance and may not be dominating. The traditional theoretical arguments mixed the above differences and thus made the whole story looks very confusing.

Based on the literature review, I would like to argue that we need to form a comprehensive and flexible view on this problem rather than dogmatic. The analyses on this problem reveal that it is probably a better way to look at the effect of inequality from analyzing a single economy for a certain period.

Reviewing the literatures, we can see that most of the recent theoretical arguments are based on modeling works (esp. endogenous growth frame). Such models are suffering four major problems: firstly, some of the key assumptions in the models are far from reality. Thus the results "proved" by these models are doubtful for looking at the real economy. This is also one of the reasons for the different results. Secondly, most of these models keep ambiguous attitudes on whether the effect of inequality on growth is dominant or not. In fact, it is very difficult to clearly illustrate this problem in modeling, that is, the relative setups usually cannot be flexible enough to incorporate all growth impetuses and all kinds of development stages into one frame. Thirdly, the

neoclassical models are pretty weak to capture the institutional factors which sometimes play a crucial role. The limitations of neoclassical frame on analyzing a complicated transitional economy like China should be noted. Finally and maybe most importantly, it is quite doubtful that the neoclassical frame is appropriate to analyze the problem since in the neoclassical frame production factors receive the revenue equal to their marginal productivity whereas the inequality problem in reality usually comes from deviation from this rule.

Concerning the empirical studies, although the fashionable time series methods are more applied recently with criticizing the cross-section and panel methods, such studies are weak to differentiate the dominant effect and partial effect and their methods are always limited by the data quality.

5.2 The implication of China's experience on this problem

The discussion of the problem focusing on one economy is very meaningful, since we argued that the effects of inequality on growth may be diversified for different development stages of different economies. China's economic growth experience during the past 30 years, for example, can lead us quickly realize that some of the above results in the literatures are obviously wrong for a transitional economy.

Firstly, as a basic fact, China's GDP growth rate has been keeping staying at a very high level when its inequality situation keeps worsening (See Figure 1.1 and Figure 1.2 in Chapter 1). An obvious implication of the above combination is that the effect of income/wealth inequality on economic growth is not dominant for China during the 30 years. Thus if we say the effect is negative, it means that inequality exerts a partially negative effect on growth and this effect is overwhelmed by other positive effects, at least it is the case for China in this early transitional period.

Further, China's experience doubts the arguments that the effect of high initial wealth/income inequality on subsequent growth is bad and will possibly lead the economy to stagnation like what Banerjee and Newman (1993), Alesina and Rodrik (1994) and Deininger and Squire (1998) discussed. As introduced in Chapter 1, China's

economic development starts from 1949 with a very equal income/wealth distribution. In the next 30 years, China's economy collapsed several times and still remained as a very poor economy in 1979. Since 1979, China's inequality situation has been keeping worsening. But the very high growth rates since then have never been really blocked. As a matter of comparison between the two periods, the economic institution matters much more than the inequality factor on growth performance.

Reviewing the literatures on the effect of income/wealth distribution on China's growth performance, most of them are empirical and they have also generated mixed results. Using panel data of provinces from 1987 to 2001, Wan, Lu and Chen (2006) estimated simultaneous systems of equations with distributed lag models. Specifically, they applied PIL (polynomial inverse lag) structure of Mitchell and Speaker (1986). With achieving a significance of 10% level, they got a nonlinear relationship and negative result irrespective of time horizons despite seemingly positive correlation between growth and inequality in post-reform China. They argued that the negative effects stem from the strong and negative influence of income inequality on physical investment which is contrary to Kaldor (1955)'s argument. But their results show that the relationship between income inequality and education level is always positive.

Wang and Cai (2006) applied the Vector Error Correction Model with Granger-causality test, discussing the effect of income inequality on investment. Based on the results, they argued that only the income level of urban residents affects the investment in both light and heavy industries with more effects on the heavy industries. The income level of rural residents only affects the light industries; also, the change of the urban income distribution affects negatively light industry but affects positively heavy industry. This means that the higher income inequality in urban areas will promote the growth of heavy industries. The paper did not discuss the overall effect for the whole economic growth but we can conclude from their arguments that the higher income inequality of both urban and rural areas will be beneficial for the investment growth in heavy industries and the effect on light industries is not clear.

Wang and Ouyang (2007) studied the effect of urban-rural income inequality with

Tale index from 1979 to 2004. The result of Tale index shows that the change of urban-rural income inequality experienced two stages' fluctuation of "V" shape: 1979-1994 and 1995-2004. Then they set up an estimation model based on the method of PVECM (Panel Vector Error Correction Model) with the year 1995 as the dummy variable which is treated as the separation line of the effects based on the observation of Tale index from 1979 to 2004. The results show that during the early period (1979-1995) the urban-rural income inequality is good for growth and during the later period (1995-2004), the disparity harms growth. They argued that their result can be explained through the theory presented by Galor and Moav (2004). Their research shows that different provinces have different level of effect although with the same sign (negative or positive) before and after 1995.

Risso and Carrera (2010) analyzed the income inequality-growth relationship in China for two periods: 1952-1978 and 1979-2007, using the Vector Error-Correction model with Johansen's cointegration technique. They achieved the results that inequality-growth relationship is positive for both periods and there is no directional causality in the second period from the Granger-causality test.

Yang and Huang (2011) used panel data of 23 Chinese provinces from 1996 to 2007 to generate the result that during the early stage of economic development when the average per capita education time below 8.973 years，income inequality will benefit economic growth and harm it during the late stage when the per capita education time above 8.973 years. They argued that since in 2007 the per capita education time is 8.185 years, till that time the income inequality had always benefited economic growth. Their results of non-linear trend is similar to that of Wang and Ouyang (2007)'s. But the separation lines of the opposite effects are completely different.

Discussions

In chapter 2, we argued from China's experience during the past 20 years that the positive effect of income/wealth inequality on growth is strongly doubtful in the long run when a transitional economy is experiencing the tertiarisation process. The channel of the negative effect is mainly from the insufficient demand for sustainable

tertiarisation.

Some of the empirical studies reviewed above support my argument on the negative effect of inequality on growth for the transitional economy in the long-run. But some of the results are pointing to the positive effect of inequality on growth. We should be more cautious about this stand. First of all, although the sign we achieved from some empirical studies is indeed positive, it doesn't mean that the inequality is good for growth from the causality point of view. Secondly, some arguments supporting the positive effect, in my view, suffer various problems. For example, the argument that the larger inequality is good for physical investment ignores the problem of insufficient demand. The argument that richer people have a higher propensity to save (Kaldor, 1955) also means that the high degree of income/wealth inequality will cause the problem of insufficient demand. But for an open economy, this problem can be partly solved by international trade before the period of tertiarisation. So the effect of inequality could be good for growth only in the short run and this is unsustainable as economy develops. As for the argument that the initial inequality is good for human capital accumulation, it is not consistent with socialism economic experience. For China, the basic education cost for people is very low which is supported by the government. Thus the inequality hasn't much relationship with human resource.

5.3 Rent-seeking space, transaction cost and economic growth

We built the positive relationship between rent distribution and income/wealth inequality in Chapter 3. The effect of rent, in theoretical sense, is not the same as the effect of inequality, although they have some similarities. The effect of rent on growth, as we will analyze, is mainly through the channel of negative externality whereas the effect of inequality on growth, as we introduced in the literature review, has its own diversified channels. However, the negative effect of inequality on economic growth in the long run can also be seen from the angle of rent-seeking. As we showed in Chapter 3, the distribution of rents is positively related to the distribution of wealth in the long run with the determining effect. If the income/wealth inequality evolution is

mainly caused by the rent distribution evolution in the long run, the effect of rent distribution on economic growth can also tell us the effect of inequality on growth.
First of all, we give the formal definition of rent-seeking space:

DEFINITION 1. The rent-seeking space is a space of the rent species R_i, rent revenue $g_i(R_i)$ and rent distribution Ω. The rent-seeking space is defined as
$\mathbb{R} =: \{(R_i, g_i(R_i), \Omega), i \in \mathbb{N}\}$

Rent-seeking space can be seen as a description of one aspect of institution. We define the rent revenue $g_i(\bullet)$ as a monotone increasing function of rent species R_i where i points to the rent specification. Rent species R_i represents the rent quantity of the specification i. The function $g_i(\bullet)$ is concave and satisfying the Inada conditions: $\lim_{R\to 0} g_i'(R_i) = \infty$ and $\lim_{R\to\infty} g_i'(R_i) = 0$. The concavity assumption means that the rent revenue is going to be larger as the rent quantity becomes larger. The increasing speed will slow down, however, when the rent goes larger. This is similar to the characteristics of the utility function and the neoclassical production function we usually apply. The rent revenue will bring utility which is the root reason behind our assumptions here. The concavity assumption here is different from the convexity assumption adopted by Murphy, Shleifer and Vishny (1993). So our analysis here is also an extension of their work. The rent distribution Ω here is defined as the distribution of rent species.
We define that the rent-seeking space $\mathbb{R}$ is a monotone increasing function of R_i, $g_i(R_i)$. The rent distribution Ω is related to R_i and rent revenue: the more equal distribution of R_i will lower down the rent revenue $g_i(R_i)$. Noting that the rent quantity is consistent with rent revenue, this is to say that the more equal rent distribution Ω automatically means R_i is lower, vice versa. Plus, the increase of rent specification i means the society is more unfair and we can see that it is consistent with more unequal distribution Ω. From the above setup and arguments, we can see that the larger rent-seeking space $\mathbb{R}$ is equivalent to the more unequal rent distribution Ω. Thus from Proposition 1 of Chapter 3, there is a positive relationship between

rent-seeking space $\mathbb{R}$ and wealth inequality in the long run. From the setup we transfer our analysis with rent distribution to the rent-seeking space.

Since the development of an economy will change the elements of rent-seeking space, the relative institutional innovations are necessary for checking the expansion of the space. Now we turn to analyze the properties of rent-seeking space:

LEMMA 1. The principle of optimal allocation of rents R_i is similar to the principle of optimal allocation of economic resources in microeconomics that the rent-owner will allocate its different rents to the point that the marginal revenue from offering each rent equal in order to maximize the total revenue of his/its whole rents.

Proof: Define a vector of rent species (R_1, R_2,$\cdots$,R_N) which is the rent endowment of a rent-owner. His goal is to maximize the total revenue from either selling the rents or using the rents themselves. Assume that the revenue function $G(R_1, R_2,\cdots,R_N)=\sum_{i=1}^{N} g_i(R_i)$ which is concave. Let R denotes the total rent he owns. We can setup an optimization problem:

Max $G(R_1, R_2,\cdots,R_N)=\sum_{i=1}^{N} g_i(R_i)$

s.t. $\sum_{i=1}^{N} R_i = R$

Solving by the Lagrangian method: $L = \sum_{i=1}^{N} g_i(R_i) + \mu(R - \sum_{i=1}^{N} R_i)$

We get $\frac{dg_1}{dR_1} = \frac{dg_2}{dR_2} = \cdots = \frac{dg_N}{dR_N} = \mu$

The second-order condition is checked which confirms the Lagrangian result is the maximum point.

This result means that the rent-owner will make each of his rent species has the same marginal revenue by allocating them rationally. The Lagrangian multiplier μ is the shadow price of each rent.

Q.E.D.

Mao (1985) discussed the principle of optimal allocation of resources with similar structure. Here we showed that the rent, as a kind of resource, is also obeying the

same allocation principle in product and resource markets. The transaction cost in the dissertation, as defined in Chapter 4, is the cost of achieving an acceptable transaction. For some cases, transactions have no relationship with rent, e.g. buying normal commodities. However, there are some transactions contain rent-seeking. The price of the rent is related to such transactions. Traditionally, transaction cost (defined by Coase) has already been embodied in the analysis on the problem of rent-seeking activities (Hartle, 1983). However, the qualitative relationship between transaction cost and rent has never been established. Following the definition of transaction cost in the dissertation, we can establish a relationship between rent and transaction cost:

LEMMA 2. The price of rent is the marginal value of the part of transaction cost which the rent is applied to reduce.

Proof: Define the price of a kind of rent (not the total rent of a rent-owner) as P_{Ri} and the transaction cost related to R_i as TC_{Ri}, we have

$$P_{Ri}=\frac{dTC_{Ri}}{dR_i} \qquad (1)$$

The intuition of this relationship should be clear: the rent is applied to reduce the transaction cost and the price of the rent thus should be equal to the marginal value of the transaction cost related to the rent.

The right side is the marginal benefit of the rent seeker and the left side is the cost. If the cost is less than the marginal benefit, the rent seeker will increase his rent demand which will eventually increase the cost and lower the marginal benefit. Oppositely, if the cost is larger than the marginal benefit, the rent seeker will reduce his rent demand which will lower the cost and increase the marginal benefit. This will lead the equalization of the two sides.

Q.E.D.

The above rule established is also consistent with the principle of demanding side of the competitive resource market. Lemma 1 and Lemma 2 point to rules of the supply

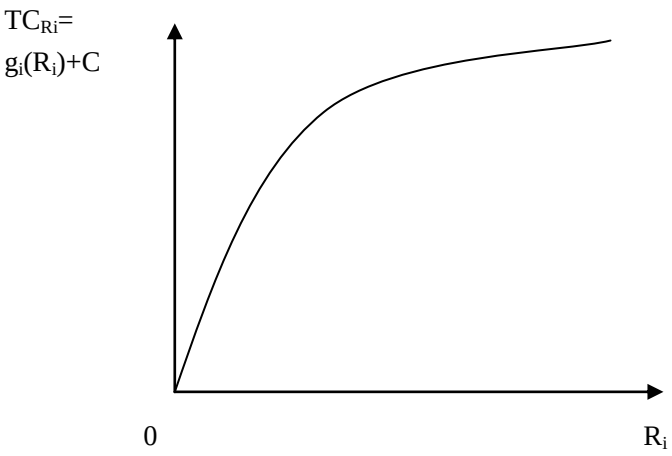
and demand sides of the rent-seeking activities. Their interacting relationship will be explored in Proposition 1.

PROPOSITION 1. The larger the rent-seeking space, the larger the total social transaction cost and the transaction cost has a concave relationship with the rent related to the transaction.

Proof: The rent-seeking space is a monotone increasing function of both rent and rent revenue. From $P_{Ri}=\frac{dTC_{Ri}}{dR_i}$ (1) we can get $TC_{Ri}=\int P_{Ri}\, dRi + C$, where C is a constant. So if P_{Ri} is larger, TC_{Ri} will be also larger. TC_{Ri} represents the transaction cost related to the rent i. Further, in equilibrium, we have the relationship between Lemma1 and Lemma 2 that $P_{Ri}=\frac{dTC_{Ri}}{dR_i} = \mu = \frac{dg_i}{dR_i}$. This means $TC_{Ri}=g_i(R_i)+C$.

The rent-seeking activity concerning rent i has both the private effect (for rent-seeker himself) and the social effect (Angeletos and Kollintzas (2000) treated them as internal effect and external effect), from this relationship we can see that the increase of rent revenue means the transaction cost associated with the rent will be larger in the private level. Since g(•) is also a concave function, we know that the transaction cost associated with the rent is also concave. This is private level transaction cost. The concavity property of transaction cost related to the rent entailed can be described by

Figure 5.1



According to the traditional rent-seeking theories, rent-seeking activity also has an effect of negative externality (social effect). Thus if R_i or $g_i(R_i)$ is larger, it will also cause the increment of social transaction cost ($TC_{\overline{Ri}}$) from the negative externality. Since $TC=TC_{Ri}+TC_{\overline{Ri}}$, the overall transaction cost of a society TC will increase. In mathematics, this means that if $dR_i/dt>0$ or $d(g_i(R_i))/dt>0$, we have $dTC/dt>0$. Thus we have the relationship that the social transaction cost TC is a monotonic increasing function of rent-seeking space.

Q.E.D.

PROPOSITION 2. The "Harberger Triangle" type negative externality of the rent-seeking space grows with a convex path as the rent-seeking space expands.

Proof: Rama (1993) adopted the concept of Harberger costs to describe the gap between the actual and potential output with restrictive regulation. The "Harberger Triangle" traditionally is used to analyze the welfare loss of monopoly or price distortions. The factor of rent-seeking can be included in the model as the monopoly degree which can be represented either by the steeper demand curve the enterprise faces or by the higher distorted price the enterprise can decide. These two directions are consistent.

Figure 5.2

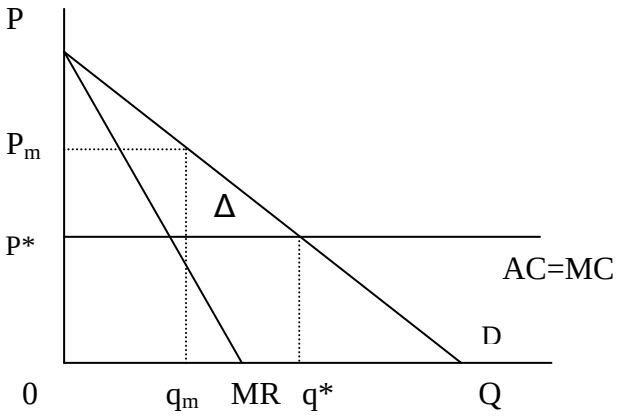


Figure 5.3

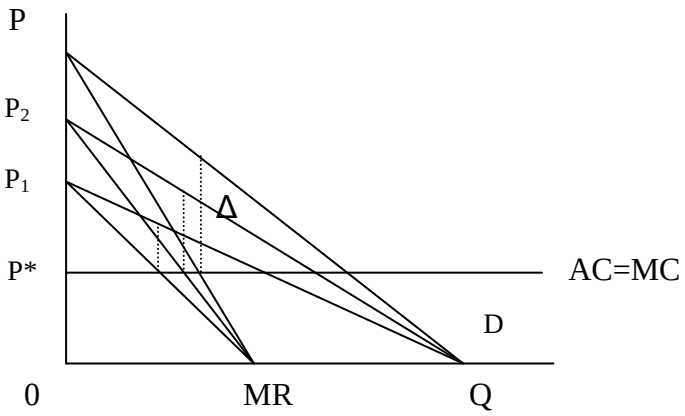


Figure 5.2 is the traditional monopoly diagram which was invented to analyze the welfare effect of rent seeking activities. The triangle Δ is usually called "Harberger Triangle" which represents the pure welfare loss caused by monopoly. The Harberger cost can be viewed as the negative externality. P* is the Pareto optimal price which represents perfect competitive equilibrium. P_m is the monopoly price which can represent the degree of rent-seeking space.

In Figure 5.3 we change the slope of the demand curve PQ to represent the change of monopoly power. P_1Q is flatter than P_2Q which is flatter than PQ. The flatter slope means the price elasticity of the demand curve D is higher. This means the market is more competitive where the rent-seeking space is smaller. Without changing the Q, we can see that the higher monopoly power owns a higher price determining position. The "Harberger Triangle" Δ is becoming smaller and smaller as the market becomes more competitive. But is this decreasing speed convex or concave?

We can check this by assuming the demand curve is $Q = \alpha - \beta_i P$ which is equal to

$P = -\frac{1}{\beta_i}Q + \frac{\alpha}{\beta_i}$. $\alpha > 0, \beta > 0$. The higher β_i means the slope is flatter. Suppose the AC curve stays at the level P*=x. The MR curve thus can be expressed as $MR = -\frac{2}{\beta_i}Q + \frac{\alpha}{\beta_i}$. With the setup, we can get the base and height expressions of the "Harberger Triangle" Δ and thus get the mathematical expression for the area of the "Harberger Triangle" Δ : $S = \frac{1}{2}\left(\frac{\alpha}{2} - \frac{\beta_i}{2}x\right)\left(\frac{\alpha}{2\beta_i} - \frac{x}{2}\right) = \frac{\alpha^2 + x^2{\beta_i}^2}{8\beta_i} - \frac{\alpha x}{4}$.

From the above formula, we have $\frac{dS}{d\beta_i} < 0$ when $x < \frac{\alpha}{\beta_i}$. Since $\frac{\alpha}{\beta_i}$ is the intercept of the demand curve, we can rule out the possibility of $x > \frac{\alpha}{\beta_i}$ since no enterprise will keep operating with no profit. This result means that the "Harberger Triangle" Δ will increase as the market becomes less competitive with larger rent-seeking space.

The second order derivative $\frac{d^2S}{d{\beta_i}^2} > 0$ which mean that the change of the area of "Harberger Triangle" Δ follows a convex path with respect to the change of β_i. This means that as the negative externality measured by "Harberger Triangle" Δ will has a convex relationship with rent-seeking space in the economy.

Q.E.D.

The rent-seeking space and the "Middle-Income Trap"

In Chapter 4 we established the results that the lower average unit transaction cost of a society (or an economy) will lead the economic size to expand in the short run and the smaller the growth rate of average unit transaction cost of a society, the higher the economic growth rate will be in the long run. From Proposition 1, it is clear that the expansion of rent-seeking space means the increase of the average unit transaction cost of the society and thus is harmful for economic growth. Transaction cost, as argued in Chapter 4, is influenced by many factors, including technology innovation, institutional innovation, human capital, etc. This means that the rent-seeking effect on the total (or unit) transaction cost is also influenced by many other factors. The rent factor alone cannot cause a determinant effect on economic growth performance. Take China for example, China's rent-seeking space has been obviously expanding sine 1980s but this factor didn't block the higher growth rate of China. Thus we know that

the role of rent only plays a relatively minor role in China's economic growth during the past 30 years, although the rent-seeing activities keep distorting the economy. This is also consistent with our previous argument that the effect of income/wealth inequality on growth is only partial.

The role of rent, however, is more crucial than we thought because of the convexity property of negative externality. Proposition 2 means that the negative effects will grow with a convex path as the rent-seeking space expands. This continuous expansion of rent-seeking space will exert more and more negative effect with a growing speed. Besides, it is argued by Tullock (1967) that the "Harberger Triangle" is only a part of the social cost of monopolies and tariffs. The opportunity cost of the resources applied in achieving the rent, for example, cannot be reflected by the triangle. This problem is called "rent dissipation" which is still open to discussion.

I would further argue that in the dynamic analysis the opportunity cost can be even much larger than that under static consideration. In Chapter 2 we argued that the rent-seeking activities blocked the development of China's tertiarisation process which is the opportunity cost during the dynamic development. Such opportunity costs are especially crucial for the long-run development of a transitional economy and such costs will become larger and larger as the development steps into the "middle-income" level.

The convex characteristics of the Harberger-type negative externality of the rent-seeking space as well as the opportunity costs argued above have an important implication. Although the effect of rent-seeking space on the transaction cost related to the rent follows a concave path, the convex growth of the negative externality and the growing opportunity cost will eventually become larger than the other growth impetuses if the rent-seeking space expansion cannot be controlled.

For a transitional economy like China, rent-seeking space controlling is crucial for its further development because the negative effects of the expansion will finally surpass all other growth impetuses and make the growth unsustainable. This discussion also confirms the results of the previous studies like Rama (1993) and Angeletos and Kollintzas (2000). And for Chinese economy, the arguments are also consistent with

the logics explained in Chapter 2.

Figure 5.4 describes the effect of continuous rent-seeking space expansion:

Figure 5.4

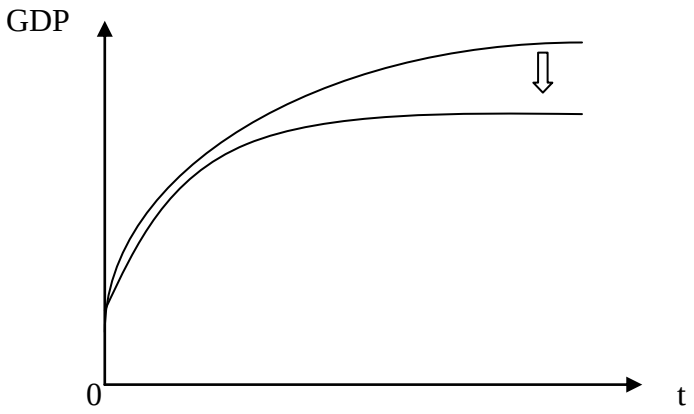


Figure 5.4 means that the economy cannot reach its potential because of the rent-seeking space expansion and the growth will slow down much earlier than its potential. For a transitional economy, this is so-called “middle-income trap” which is often faced by development economies. The background of it, based on the analysis here, is the institutional problems concerning rent-control. So institutional innovation concerning rent is a task a transitional economy must resolve and this is more difficult than the ways neoclassical growth theories suggested, like technology innovation. This can also explain the puzzle that the “should-be” quick convergence argued by exogenous growth theories usually fails in reality.

Discussions

The above analysis based on rent shows that the rent-seeking space expansion is harmful for economic growth and will block the growth in the long run if the expansion is out of control. Combining the arguments at the beginning of the section that we established the positive relationship between higher income/wealth inequality and larger rent-seeking space in the long run, the arguments here also confirms that income/wealth inequality has a negative effect on economic growth in the long run, if

the income/wealth inequality evolution is mainly caused by the rent distribution evolution in the long run. This is also consistent with the results of Model 2 in Chapter 3.

5.4 Concluding remarks

The chapter discussed the effect of income/wealth inequality on economic growth with exploring the properties of rent-seeking space. Our research in this chapter is related to previous chapters. We built a relationship among rent-seeking space, income/wealth inequality and economic growth based on the frame built in Chapter 4. The main result is that rent-lead inequality exerts a negative effect on economic growth in the long run.

As an implication, rent-control will be good for both equality and growth. This requires institutional updates during growth. The successful growth story of China during the past 30 years can be seen as a "take-off" model for a transitional economy. The future success from a middle-income economy to a developed economy needs relative institutional reforms concerning rent-seeking space. This will be an even bigger challenge for China.

Chapter 6

Conclusions and Implications of the Dissertation

Prosperity and common prosperity are always two different issues. However, they are closely related to each other. This dissertation studied the relationship between economic growth and income/wealth distribution dynamically. The background of this research is a huge transitional economy: mainland China. Our analysis reveals that institutional arrangements play a crucial role in this relationship.

As a transitional economy, China is an ideal case for institutional analysis. The political and economic institutional arrangements of China are obviously taking crucial effects on its economic development. The dissertation tries to present a comprehensive and incisive economic analysis on this issue which can be summarized as several conclusions and implications for transitional economies' development:

1. If a developing economy hopes to successfully update itself into a developed economy, the transitional process needs the cooperation of institutional innovations. The main point is to limit the rent-seeking space through continuous in-time institutional innovations, especially during the stage of tertiarisation process. Both political and economic institutions matter for the transition. Mere policy innovations are not sufficient.

2. Different periods during the transition have different growth motivations. A successful growth pattern during the early stage doesn't mean it will be successful for the later stage.

3. For a transitional economy adopting double-track economic system like China, the process of combining the double-track into a single track is a must for developing the

marketing economy. The political institution matters during the process through its effect on the rent-seeking space. A higher quality of law-ruled society is needed for further advancement.

4. The economic orders behind the neoclassical frame, as argued in Chapter 3, are good for income/wealth equality with continuous economic growth. The key points are to equally distribute the rent-seeking space and to realize the perfect competitive market.

The dissertation also sheds some light on economic theories:

1. The frame of neoclassical economics cannot be brought to analyze transitional economy without any change on the relative models.

2. The new institutional economics has a very strong explaining power on transitional economic development. The marginal analysis can be used to analyze institutional factors and is useful on exploring relative problems. The combination of neoclassical and new institutional economics can generate much bigger explaining power on economic reality.

3. Transaction cost (defined in the dissertation) is the motivation of economic growth which is the channel for all growth-enhance factors like technology or human capital.

4. Inequality only partially affects the economic growth performance. This point should be paid attention for future modeling works. But this partial effect can change into a dominant effect especially for the transitional economy experiencing tertiarisation.

5. The angle of equilibrium transaction speed reveals some new ideas on the issue of economic crisis. We pointed out that the difference between early and later stages

during crisis should be paid attention to.

6. The crucial way to create more equal income/wealth distribution in the long run is to control the rent-seeking space in the way to make the rent distribution more equal. This is the obligation of politics/government during economic development.

A debate with Karl Marx: does the economic base determine the superstructure?

Marxism theory has a very famous proposition that the base determines the superstructure. The base and superstructure are the two parts forming human society. The base is related to economic production as what Marx said in 1859:

In the social production of their existence, men inevitably enter into definite relations, which are independent of their will, namely the relations of production appropriate to a given stage in the development of their material forces of production. The totality of these relations of production constitutes the economic structure of society, the real foundation, on which arises a legal and political superstructure.

Source: Marx, K. (1859) *A Contribution to the Critique of Political Economy*, Moscow: Progress Publishers: Notes by R. Rojas.

The superstructure of a society includes culture, institutions, political structure, rituals, etc. Such contents can be included in a general and broader definition of institution. Kasper and Streit (1998) divided the definition of institution into two parts: internal institution and external institution. The internal institution is reflected by customs, ethical norms, etc. which evolve from human experience that tend to serve people the best. The external institutions which are often reflected as law regulations are imposed and enforced from above like agents who are authorized by a political process.

The case of China's development since reform and opening up, as analyzed in the

dissertation, shows the opposite direction of the base-superstructure nexus. It is the superstructure that determines the economic base which further influences the production force. The superstructure didn't change much as the base changes, although the voice asking for change has never disappeared. The economic base's relationship with production force is also difficult to be argued as that the production relation is determined by production forces. The production relationship, as analyzed in the dissertation, is prone to be determined by superstructure (institutions).

Of course, the dissertation also shows that the long-run relationship and the short-run relationship can be opposite. The argument of Marx on the relationship between the base and superstructure is prone to be a long-run historical view. This view, however, needs to be adjusted when we look at the relative short-run history.

Notes

1. The household responsibility system (HRS) was created by several villagers in Xiaogang Village in Fengyang County, Anhui Province, China. It was the end of 1978. As a try to promote the production, they took responsibility for their own gains and losses, with an agreement that if any of them were to go to jail for secretly embarking on this illegal system, the others must take care of their children. The test generated incredible results that at the end of 1979, the one-year production is equal to the 5 years' total production from 1966 to 1970. Seeing this result, the Central Rural Work Conference at the end of 1979 decided that the poorest residents in rural areas would be allowed to engage in this system. At the end of 1980, 14% of the production teams around the country followed the system and all production teams under the system achieved remarkable results that year. So in 1981 the government started to promote the system across the country. By the end of the year, 45% of production teams were in the system. Till 1984, 99% of the production teams were included. This story is usually seen as the beginning of China's reform and opening up.

2. Proof: Let $\frac{r_i-\rho}{\theta} - r_i = \eta$, if $\eta>0$, when $t\to\infty$, $e^{\eta t}\to\infty$, thus the possibility is excluded; if $\eta<0$, from (3) we have $1 - e^{\eta t} \le -\eta$. When $t\to\infty$, $e^{\eta t}\to 0$. This means that $1 \le -\eta$ which is our boundedness condition: $\eta \le -1$. Q.E.D.

3. The result can be seen from that the second term can control the last two terms and the first term can control the 3rd term, when t is sufficiently large.

4. Domar (1947)'s logic is as follows: Suppose the economy starts from full employment with total income 150 billion dollars. 10% is used for saving. With condition S=I, we have 15 billion dollars invested. But this additional supply needs equivalent demand power increase. But keeping income constant with 150 billion, the real income increase can only fulfilled by deflation, which Domar offered 4 reasons to

argue not always possible, including monopolistic power, deflation's bad effects and relative price reduction speed between consumers' goods and investment goods. So with assuming a constant general price level, the investment needs to expand every year to keep full employment, that is, the investment next year I_{t+1} must be greater than S_t. This means that if the investment next year I_{t+1} be equal to S_t, there will be unemployment in period t+1.

References

Arrow, K. J. (1962) The Economic Implication of Learning by Doing, *The Review of Economic Studies*, 29, pp. 155-173.

Alesina, A. and Perotti, R. (1993) *The Political Economy of Growth: A Critical Survey*, Mimeo. (Harvard University, Cambridge, MA)

Alesina, A. and Rodrik, D. (1994) Distributive politics and economic growth, *The Quarterly Journal of Economics*, 109, pp. 465-490.

Alesina, A. and Perotti, R. (1996) Income Distribution, Political Instability, and Investment, *European Economic Review*, 40, pp. 1203-1228.

Aghion, P. and Bolton, P. (1997) A Theory of Trickle-Down Growth and Development, *The Review of Economic Studies*, 64, pp. 151-172.

Aghion, P., Caroli, E. and Garcia-Penalosa, C. (1999) Inequality and Economic Growth: The Perspective of the New Growth Theories, *Journal of Economic Literature*, 37, pp. 1615-1660.

Angeletos, G. M. and Kollintzas, T. (2000) Rent Seeking/Corruption and Growth: A Simple Model, *CEPR Discussion Paper* No. 2464.

Andrade, J.A.S., Duarte, A.P.S. and Simoes, M.C.N. (2011) Inequality and Growth in Portugal: a time series analysis, *GEMF Working Papers,* Universidade de Coimbra.

Acemoglu, D. and Robinson, J.A. (2012) *Why Nations Fail: The Origins of Power, Prosperity and Poverty*, London: PROFILE BOOKS LTD.

Bronfenbrenner, M. (1971) *Income Distribution Theory*, Chicago: Aldine-Atherton.

Bourguignon, F. (1981) Pareto Superiority of Unegalitarian Equilibria in Stiglitz's of Wealthy Distribution with Convex Saving Function, *Econometrica*, 49, pp. 1469-1475.

Banerjee, A. V. and Newman, A. F. (1993), Occupational Choice and the Process of Development, *Journal of Political Economy*, 101, pp. 274-298.

Birdsall, N., Ross, D. and Sabot, R. (1995) Inequality and Growth Reconsidered: Lessons from East Asia, *The World Bank Economic Review*, 9, pp. 477-508.

Barro, R. (2000) Inequality and Growth in a Panel of Countries, *Journal of Economic*

Growth, 5, pp. 5-32.

Banister, J. (2007) Manufacturing in China Today: Employment and Labor Compensation, Conference Board Economics Program Working Paper 07-01.

Commons, J. R. (1934) *Institutional Economics*, Madison: University of Wisconsin Press.

Coase, R. (1937) The Nature of the Firm, *Economica*, 4, pp.386-405.

Coase, R. (1960) The Problem of Social Cost, *Journal of Law and Economics*, 3, pp. 1-44.

Chenery, H. B. (1982) Industrial and Growth: The Experience of Large Countries, *World Bank Staff Working Papers*, No. 539.

Chatterjee, S. (1994) Transitional Dynamics and the Distribution of Wealth in a Neoclassical Growth Model, *Journal of Public Economics*, 54, pp. 97-119.

Cheung, S.N.S. (1998) The Transaction Costs Paradigm 1998 Presidential Address Western Economic Association, *Economic Enquiry*, 36, pp. 514-521.

Caselli, F. and Ventura, J. (2000) A Representative Consumer Theory of Distribution, *The American Economic Review*, 90, pp. 909-926.

Cheng, Y. H. (2007) The Evolution of China Gini Coefficients and Its Urban-Rural Constructions since Reform and Opening Up (改革开放以来中国总体基尼系数的演变及其城乡分解), *China Social Science* (in Chinese), 4, pp.45-60.

Chen, J. D. and Dai, D. (2011) Promoting the process of urbanization and reduce the income inequality of Chinese people (加快城镇化进程与改善我国居民的收入不平等), *Public Finance Research* (in Chinese), pp. 48-52.

Domar, E. D. (1947) Expansion and Employment, *The American Economic Review*, 37, pp. 34-55.

Deninger, K. and Squire, L. (1998) New Ways of Looking at Old Issues: Inequality and Growth, *Journal of Development Economics*, 57, pp.259-287.

Dollar, D. and Kraay, A. (2002) Growth is good for the Poor, *Journal of Economic Growth*, 7, pp. 195-225.

Dollar, D. and Kraay, A. (2002) *Growth is good for the Poor: A Response to Lubker, Smith and Weeks*, World Bank.

Forbes, K. J. (2000) A Reassessment of the Relationship between Inequality and Growth, *The American Economic Review*, 90, pp. 869-887.

Feng, J. L. (2004) Income Disparity in China: A Review of Chinese Studies, *Income Disparities in China An OECD Perspective*, OECD.

General Investigation Team of National Bureau of Statistics, P. R. China (2004) The Development of Private Economy, *Reference on Economic Research (经济研究参考)*, 22.

Galor, O. and Zeira, J. (1993) Income Distribution and Macroeconomics, *The Review of Economic Studies*, 60, pp. 35-52.

Galor, O. and Tsiddon, D. (1997a) Technology Progress, Mobility, and Economic Growth, *The American Economic Review*, 87, pp. 363-382.

Galor, O. and Tsiddon, D. (1997b) The Distribution of Human Capital and Economic Growth, *Journal of Economic Growth*, 2, pp. 93-124.

Galor, O. and Moav, O. (2004) From Physical to Human Capital Accumulation: Inequality and the Process of Development, *The Review of Economic Studies*, 71, pp. 1001-1026.

Gobbin, N. and Rayp, G. (2008) Different ways of looking at old issues: a time-series approach to inequality and growth, *Applied Economics*, 40, pp.885-895.

Gerth, K. (2010) *As China goes, so goes the world: how Chinese consumers are transforming everything*, 1st Paperback Edition, New York: Hill and Wang.

Guo, Q. W. and Jia, J. X. (2005) Estimating Total Factor Productivity in China, *Economic Research Journal* (in Chinese), 6, pp. 51-60.

Harrod, R. F. (1939) An Essay in Dynamic Theory, *The Economic Journal*, 49, pp.14-33.

Hartle, D. G. (1983) The Theory of "Rent Seeking": Some Reflections, *The Canadian Journal of Economics*, 16, pp. 539-554.

Hodgson, G. M. (1988) *Economics and Institutions: A Manifesto for a Modern Institutional Economics*, Cambridge (UK): Polity Press.

Hsieh, C. T. and Klenow, P. J. (2009) Misallocation and Manufacturing TFP in China and India, *The Quarterly Journal of Economics*,124, pp. 1403-1448.

IMF (2010) *World Economic Outlook: Recovery, Risk, and Rebalancing*, October.

Jarque, C. M. and Bera, A. K. (1987) A Test for Normality of Observations and Regression Residuals, *International Statistical Review*, 55, pp. 163-172.

Kaldor, N. (1955) Alternative Theories of Distribution, *The Review of Economic Studies*, 23, pp. 83-100.

Kuznets, S. (1955) Economic Growth and Income Inequality, *The American Economic Review*, 45, pp. 1-28.

Kaldor, N. (1957) A Model of Economic Growth, *The Economic Journal*, 67, pp. 591-624.

Krueger, A. O. (1974) The Political Economy of the Rent-Seeking Society, *The American Economic Review*, 64, pp.291-303.

Kasper, W. and Streit, M. (1998) *Institutional Economics: Social Order and Public Policy*, UK: Edward Elgar Publishing Ltd.

Kuijs, L. and Wang, T. (2005) China's Pattern of Growth: Moving to Sustainability and Reducing Inequality, *World Bank Policy Research Working Paper*, No. 3767.

Lucas, R. E. Jr. (1992) On Efficiency and Distribution, *The Economic Journal*, 102, pp. 233-247.

Lin, Y. F., Cai F. and Li, Z. (1994) *The China's Miracle: Development Strategy and Economic Reform* (in Chinese), Shanghai People's Publishing House.

Lau, L. J., Qian, Y. Y. and Roland, G. (1997) Pareto-Improving Economic Reforms through Dual-Track Liberalization, *Economics Letters*, 55, pp. 285-292.

Lin, Y. F. and Tan, G. F. (1999) Policy Burdens, Accountability and Soft Budget Constraint, *The American Economic Review*, 89, pp. 426-431.

Lau, L. J., Qian, Y. Y. and Roland, G. (2000) Reform without Losers: An Interpretation of China's Dual-Track Approach to Transition, *Journal of Political Economy*, 108, pp. 120-143.

Liu, X. X. (2000) The Influence of Ownership Structure of China's Industrial Enterprises on Efficiency Difference—An Empirical Study Based on 1995 National Survey Data on Industrial Enterprises，*Economic Research Journal* (in Chinese), 2, pp. 17-25.

Li, H. Y., Xie, D. Y. And Zou, H. F. (2000) Dynamics of Income Distribution, *The Canadian Journal of Economics*, 33, pp. 937-961.

Lin, Y. F. and Li, Y. J. (2001) Promoting the Growth of Medium and Small Sized Enterprises through the Development of Medium and Small Sized Financial Institutions, *Economic Research Journal* (in Chinese), 1, pp. 10-18.

Lubker, M., Smith G. and Weeks J. (2002) Growth and the Poor: A Comment on Dollar and Kraay, *Journal of International Development*, 14, pp. 555-571.

Lin, Y. F. and Liu, P. L. (2003) China Economic Development Strategy and Regional Income Disparity, *Economic Research Journal* (in Chinese), 3.

Lu, M. (2003) Why Does Reform Have not Improved the Relative Labor Productivity of China's SOEs? *China Economic Quarterly*, 2, pp. 833-856.

Lucas, R. E. Jr. (2004) Life Earnings and Rural-Urban Migration, *Journal of Political Economy*, 112, pp. S29-S59.

Lin, Y. F. (2011a) *Demystifying the Chinese Economy*, 1st Edition, Cambridge University Press.

Lin, Y. F. (2011b) *China and the Global Economy*, Remarks at the Conference "Asia's Role in the Post-Crisis Global Economy" San Francisco Federal Reserve Bank, November 29th.

Lin, Y. F. (2012) *New Structure Economics: a Framework of Rethinking Development*, The World Bank.

Marx, K. (1859) *A Contribution to the Critique of Political Economy*, Moscow: Progress Publishers: Notes by R. Rojas.

Marx, K. (1885) *Capital*, *The 2nd Volume*, 1975 Chinese Edition, Beijing: The People's Press, pp. 435-592.

Mirrlees，J. A. (1971) An Exploration in the Theory of Optimum Income Taxation, *The Review of Economic Studies*, 38, pp. 175-208.

Mao, Y. S. (1985) *Principle of Optimal Allocation: Mathematical Foundation of Economics*, 1st Edition, Chengdu: Sichuan People's Press.

Mitchell, D. W. and Speaker, P. J. (1986) A Simple, Flexible Distributed Lag Technique: The Polynomial Inverse Lag, *Journal of Econometrics*, 31, pp. 329-340.

Mankiw, G., Romer, P. and Weil, D. (1992) A Contribution to the Empirics of Economic Growth, *The Quarterly Journal of Economics*, 107, pp.407-437.

Murphy, K. M., Shleifer, A. and Vishny, R. W. (1993) Why is Rent-Seeking So Costly to Growth? *The American Economic Review*, 83, pp. 409-414.

Mas-colell, A., Whinston, M. D. and Green, J. R. (1995) *Microeconomic Theory*, Chapter 4, New York: Oxford University Press.

Mah, J.S. (2003) A Note on Globalization and Income Distribution—The Case of Korea, 1975-1995, *Journal of Asian Economics*, 14, pp. 157-164.

Maddison, A. (2003) *The World Economy: Historical Statistics*, OECD Development Centre.

Maddison, A. (2006) *The World Economy*, OECD Development Centre.

Maddison, A. (2007) *Chinese Economic Performance in the Long Run, 960-2030AD*, 2nd Edition, Shanghai: People's Press Chinese Edition.

Mao Y. S. (2008) Seeking the Answer to Fast Economic Growth from History, *Mao Yushi Talks about Economics: from Micro to Macro*, China Shanxi Publishing Ltd.

McMillan, M.S. and Rodrik, D. (2011) Globalization, Structural Change and Productivity Growth，*NBER Working Paper*, No. 17143.

Management School of Beijing Normal University and Governmental Management Graduate School (2011) *2011 China Development Report of People's Well-being*, Beijing Normal University Press.

Meng, X. (2012) Labor Market Outcomes and Reforms in China, *Journal of Economic Perspective*, 26, pp. 75-102.

North, D. C. and Thomas, R. P. (1973) *The Rise of the Western World: A New Economic History*, Cambridge: the University Press.

North, D. C. (1987) Institutions, Transaction Costs and Economic Growth, *Economic Inquiry*, 25, pp. 419-428.

North, D. C. (1989) Institutions and Economic Growth: A Historical Introduction, *World Development*, 17, pp.1319-1332.

North, D. C. (2002) *The Economy in Change and Economics*, Speech made in Peking University China Centre for Economic Research.

National School of Development, Peking University (2013) China in the Future 10 Years. *CCER Working Paper* No. R201301.

Okun, A. M. (1975) *Equality and Efficiency: The Big Tradeoff*, Washington, D.C.: The Brookings Institution.

OECD (2011) *Growing Income Inequality in OECD Countries: What Drives it and How Can Policy Tackle it?* OECD Forum on Tackling Inequality, Paris.

Perotti, R. (1994) Income Distribution and Investment, *European Economic Review*, 38, pp. 827-835.

Persson, T. and Tabellini, G. (1994) Is Inequality Harmful for Growth? *The American Economic Review*, 84, pp. 600-621.

Parente, S. and Prescott, E. (1994) Barriers to Technology Adoption and Development, *Journal of Political Economy*, 102, pp. 298-321.

Parente, S. (2001) The Failure of Endogenous Growth, *Knowledge, Technology and Policy*, 13, pp.49-58.

Peneder, M. (2002) Industrial Structure and Aggregate Growth, *WIFO Working Papers*, No. 182.

Patridge M. D. (2005) Does Income Distribution Affect U.S. State Economic Growth? *Journal of Regional Science*, 45, pp. 363-394.

Ramsey, F. P. (1928) A Mathematical Theory of Saving, *The Economic Journal*, 38, pp.543-559.

Romer, P. (1986) Increasing Returns and Long-Run Growth, *Journal of Political Economy*, 94, pp. 1002-1007.

Romer, P. (1990) Endogenous Technological Change, *Journal of Political Economy*, 98, pp. 71-102.

Rebelo, S. (1991) Long-Run Policy Analysis and Long-Run Growth, *Journal of Political Economy*, 99, pp. 500-521.

Rama, M. (1993) Rent Seeking and Economic Growth: A Theoretical Model and Some Empirical Evidence, *Journal of Development Economics*, 42, pp.35-50.

Ravallion, M. (2001) Growth, Inequality and Poverty: Looking Beyond Averages, *World Development*, 29, pp. 1803-1815.

Risso, W. A. and Carrera, E.J.S. (2010) Inequality and Economic Growth in China: pre and post-reform periods, UNIVERSITÀ DEGLI STUDI DI SIENA, QUADERNI DEL DIPARTIMENTO DI ECONOMIA POLITICA No.600.

Smith, A. (1776) *The Wealth of Nations*, 2000 Edition, New York: The Modern Library.

Solow, R. M. (1956) A Contribution to the Theory of Economic Growth, *The Quarterly Journal of Economics*, 70, pp. 7065-7094.

Sraffa, P. (1960) *Production of Commodities by Means of Commodities: Prelude to a Critique of Economic Theory*, Cambridge: Cambridge University Press.

Stiglitz, J. (1969) Distribution of Income and Wealth among Individuals, *Econometrica*, 37, pp. 382-397.

Streeck, W. (1992) *Social Institutions and Economic Performance: Industrial Relations in Advanced Capitalist Economies*, London: SAGE Publication Ltd.

Schefold, B. (1997) *Normal Prices, Technical Change, and Accumulation*, New York: St. Martin's Press.

Siebert, H. (1998) Commentary: Economic Consequences of Income Inequality, *Symposium of the Federal Reserve Bank of Kansas City on Income Inequality: Issues and Policy Options*, pp. 265-281.

Song, Z., Storesletten, K. and Zilibotti, F. (2011) Growing like China, *The American Economic Review*, 101, pp. 202-241.

Stiglitz, J. (2012) *The Price of Inequality: How Today's Divided Society Endangers Our Future*, 1st Edition, New York: W. W. Norton & Company.

Tullock, G. (1967) The Welfare Costs of Tariffs, Monopolies, and Theft, *Western Economic Journal (now Economic Inquiry)*, 5, pp. 224-232.

Thurow, L. C. (1970) Analyzing the American Income Distribution, *The American Economic Review*, 60, pp. 261-269.

Tinbergen, J. (1975) *Income Distribution: Analysis and Policies*, Amsterdam: North-Holland.

Tollison, R. D. (1982) Rent Seeking: A Survey, *KYKLOS*, 35, pp. 575-602.

Williamson, O. E. (1975) *Markets and Hierarchies: Analysis and Antitrust*

Implications, New York: Free Press.

Williamson, O. E. (1985) *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*, New York: The Free Press.

Williamson, O. E. (1995) *The Mechanism of Governance*, New York: Oxford University Press.

Wu, J. (2000) China's Economic Reform: Past, Present and Future, *Perspectives*, 1.

Wang, X. L. and Fan, G. (2004) Analysis on the Regional Disparity in China and the Influential Factors, *Economic Research Journal* (in Chinese), 1, pp. 33-44.

Wu, X. M. and Perloff, J. M. (2005) China's Income Distribution, 1985-2001, *The Review of Economics and Statistics*, 87, pp. 763-775.

Wan, G. H, Lu, M. and Chen, Z. (2006) The Inequality-Growth Nexus in the Short and Long Run: Empirical Evidence from China, *Journal of Comparative Economics*, 34, pp. 654-667.

Wang, T. S. and Cai, Y. Z. (2006) Impact of Income Distribution on Capital Accumulation and Investment Structure: An Empirical Study of China since the Reform and Opening-up, *China Social Sciences* (in Chinese), 1. pp.4-14.

Wang, X. L. (2007) Grey Revenue and Inhabitants Income Inequality, *China Taxation* (中国税务), 10.

Wang, S. P. and Ouyang, Z. G. (2007) The Rural-urban Income Disparity and Its Effects to Economic Growth in the Case of China, *Economic Research Journal* (in Chinese), 10, pp.44-55.

Wooldridge, J. M. (2009) *Introductory Econometrics: A Modern Approach*, 4th Edition, International Student Edition, SOUTH-WESTERN CENGAGE Learning.

Wang, Y. F. (2010) The Estimation of Income Distribution in Urban and Rural China: 1985-2009, unpublished working paper, CCER, School of National Development, Peking University.

Xu, D.Q. and Li, X. (2011) *Understanding the Gap between the Rich and the Poor in China* (in Chinese), 1st Edition, Beijing: Machine Industry Press.

Yang, X. K. and Borland, J. (1991) A Microeconomic Mechanism for Economic Growth, *Journal of Political Economy*, 99, pp. 460-482.

Yang, J. and Huang, X. (2011) *Inequality and Growth* (in Chinese), 1st Edition, Beijing: Science Press.

Zhao, R. W., Li, S. and Riskin, C. (1999) *Re-study of Income Distribution Among Residents in China*, China Financial and Economic Press.

Zhu, X. D. (2012) Understanding China's Growth: Past, Present, and Future, *Journal of Economic Perspectives*, 26, pp. 103-124.