

Regional Inequality in Guangdong Province

Volume 1

Hong Yu

A Thesis Submitted for the Degree of Doctor of Philosophy

**School of East Asian Studies
The University of Sheffield**

August 2008

To my dear parents

Abstract

Guangdong has made impressive achievements in economic growth during the reform period and has been one of the fast growing provinces in China since 1978. However, regional inequality of income in Guangdong has intensified during the same period. A huge income gap between the Pearl River Delta Area (PRDA) and the peripheral regions has been found in this province. The Gini coefficient and coefficient of variation (CV) for per capita income in Guangdong's counties had been showing a steady upward trend between 1980 and 2005, not only in terms of the absolute figures of CV and Gini coefficient, but also the dramatic upward trend. The benefits of rapid economic growth have not been distributed equally among the different regions in Guangdong. The economic growth in the peripheral, hill regions has been much slower than in the PRDA. As far as income per capita is concerned, the striking regional difference between the PRDA and its periphery is evident.

Based on field research interview, analysis of geographic characteristics and regression model analysis, this paper attempts to investigate the main causes of regional disparity in Guangdong. The findings suggest that two inherited locational advantages of coast and market proximity have made a significant contribution to unbalanced regional development since 1980. In addition, the agglomeration effect is crucial in understanding the core – periphery pattern of regional inequality. This study argues that the spatial economic inequality within Guangdong is mainly caused by the double-strengthening mechanism of both geographical location and agglomeration effects. During the reform period, this mechanism has been the key factor in shaping spatial economic inequality and it is likely to cause further widening of such inequality within Guangdong in the future.

Preface and Acknowledgements

Regional disparity makes a difference to economic growth and worsening spatial income disparities require effective resolution. Regional development inequality is a real threat to China in that long-term economic growth becomes difficult to achieve. Worsening regional inequality and long-term economic development will not stand side by side and reduction of this disparity is the key to dealing with many other serious challenges that face China, including poverty and underdevelopment of health and education services.

Guangdong takes the lead in many aspects of economic development. It was also one of the earliest regions in China to implement the 'reform and open-door' policies. How should the people judge these reform policies? By focusing case study on Guangdong, this research attempts to offer meaningful and illuminating discussion of the serious issue of regional inequality in China and the re-consideration of 'reform and open-door' policies on economic growth and human development which could add to the knowledge of policymakers and scholars.

Guangdong province has made dramatic advancements in economic growth since 1978. However, the core-periphery pattern of regional inequality within Guangdong has shown an upward trend and the imbalance is severe. I have two main objectives in conducting this research. First, by adopting time-series inequality measurement, I will systematically analyze intra-regional development disparities in terms of economic and human aspects within Guangdong during the reform period. Secondly, I intend to provide a conclusive judgment on the main cause of widening regional inequality in this southern pioneer province.

This research comprises three main parts. The first section presents a literature review of approaches to measurement of regional inequality, whilst traditional economic geography, economic convergence and divergence theories are also assessed here. The second and third parts discuss measurement of regional inequality within Guangdong and analyze the main factors causing the worsening regional inequality within Guangdong, respectively.

First, I would like to take this opportunity to express thanks to my dear parents for their sacrifice and support during my PhD research. I am deeply indebted to my PhD supervisor, Prof. Chris Bramall. He has not only spent immeasurable amounts of time in giving me professional advice and valuable comments, but also provided encouragement throughout my research. I would also express thanks to the Provincial Government of Guangdong and other relevant administration organizations for offering me interview opportunities during my field research. In addition, my grateful thanks go to the staff in the English Language Teaching Center, University of Sheffield for their valuable language support and academic assistance throughout my PhD thesis writing. I also acknowledge the technical assistance given by Ms Chunqu Xu and other members of the South China Agricultural University and Mr Chaopeng Guo at the Department of Information Studies, University of Sheffield. Moreover, I would like to express my sincere thanks to the Universities' China Committee in London (UCCL). Without the research funding provided by the UCCL, I would have been unable to conduct the field research in Guangdong.

This research has certain limitations and omissions and mistakes will inevitably occur. I accept full responsibility for any misinterpretation of scholars' theories and arguments.

Hong Yu

Sheffield
October 2008

TABLE OF CONTENTS

LIST OF TABLES	10
LIST OF FIGURES	12
LIST OF MAPS	13
CHAPTER 1 GENERAL INTRODUCTION	15
1.1 The Issue of Regional Inequality	15
1.2 The Significance of Guangdong	20
1.3 Research Structure Overview	23
CHAPTER 2 LITERATURE REVIEW PART I: INEQUALITY MEASUREMENT APPROACHES	26
2.1 Introduction	26
2.2 Inequality Measurement Approaches: GDP and HDI	26
2.2.1 GDP Measurement Approach: Advantages and Criticism	26
2.2.2 HDI Measurement Approach: Advantages and Criticisms	30
2.2.3 Comparison and Evaluation: GDP and HDI	35
2.3 Happiness Theory	55
2.4 Conclusion	64
CHAPTER 3 LITERATURE REVIEW PART II: TRADITIONAL ECONOMIC GEOGRAPHY	66
3.1 Introduction	66
3.2 Traditional Economic Geography	66
3.2.1 Transportation Costs	66
3.2.2 Locational Advantage: Proximity to the Coast or Ocean-Navigable Rivers	68
3.2.3 Geographical Advantage: Market Proximity	78
3.2.4 Infrastructure, Transportation Costs and Economic Development	81
3.2.5 The Effect of Railroad on Economic Growth	82
3.3 Traditional Economic Geography: Its Critics	87
3.3.1 Criticism of Transportation Improvement	87
3.3.2 Cairncross's Theory: The Death of Distance	91
3.3.3 Criticism of the Location Advantage	95
3.4 Conclusion	98
CHAPTER 4 LITERATURE REVIEW PART III: REGIONAL ECONOMIC DIVERGENCE	100
4.1 Introduction	100

4.2 New Economic Geography	100
4.2.1 The Agglomeration Effect and Economic Growth	100
4.2.2 History and Agglomeration	103
4.2.3 Factor Concentration, Agglomeration and Economic Growth	105
4.2.4 Opposing Forces to Agglomeration	115
4.2.5 The Self-Reinforcing Pattern of Agglomeration	118
4.2.6 New Economic Geography: its Critics	122
4.2.7 Conclusion	128
4.3 Other Theories on Regional Divergence	129
4.3.1 Myrdal's Theory: Spread and Backwash Effects	129
4.3.2 Friedmann's Theory: Regional Economic Polarization	132
4.3.3 Hirschman's Theory: the Growth Pole	134
4.3.4 Endogenous Growth Theory	135
4.4 Final Conclusion	137
CHAPTER 5 LITERATURE REVIEW PART IV: REGIONAL ECONOMIC CONVERGENCE	140
5.1 Introduction	140
5.2 Regional Economic Convergence	140
5.2.1 Kuznets's Hypothesis: the Inverted U-curve of Regional Inequality	140
5.2.2 Neoclassical Growth Theory	143
5.3 World Evidence Discussion: Regional Inequality Tendency	150
5.3.1 Evidence against Kuznets's Hypothesis	150
5.3.2 Evidence against Myrdal's Spread Effect Thought'	151
5.3.3 Evidence against Neoclassical Growth Theory	153
5.4 Conclusion	161
CHAPTER 6 GEOGRAPHY, ECONOMIC GROWTH AND REGIONAL INEQUALITY IN CHINA	163
6.1 Introduction	163
6.2 China and Traditional Economic Geography	163
6.2.1 Geographical Advantage and Coastal Proximity	163
6.2.2 Transportation Costs and Economic Development	166
6.3 Industrial Cluster and Economic Agglomeration in China	172
6.3.1 History and Agglomeration	174
6.4 The Developing Trend of Regional Inequality	179
6.5 Skinner's Macroregion Model	182
CHAPTER 7 METHODOLOGY AND FIELDWORK DESIGN	187
7.1 Introduction	187
7.2 Research Methodology and Fieldwork Study Design	187
7.2.1 Research Methods: Questionnaire and Interview	187
7.2.2 Research Sample	194
7.2.3 Research Reliability and Validity	199

7.2.4 Ethical Issues	205
7.3 Conclusion	206
CHAPTER 8 CHINESE DATA, AND ASSOCIATED LIMITATIONS	208
8.1 Introduction	208
8.2 Analysis of Official Statistical Data in China	208
8.2.1 The Quality of Chinese GDP Statistical Data: Its Critics	208
8.2.2 The 2004 China Economic Census	213
8.2.3 The Quality of Chinese GDP Statistics	218
8.2.4 Official Statistical Data of Guangdong: Reliability and Accuracy	225
8.3 Conclusion	229
CHAPTER 9 MEASUREMENT OF REGIONAL INEQUALITY WITHIN GUANGDONG	231
9.1 Introduction	231
9.2 Measurement of Regional Economic Inequality	231
9.2.1 Inter-regional Economic Disparities	231
9.2.2 Inter-County Income Inequality within Guangdong	240
9.3 Measurement of Regional Income Inequality within Guangdong	259
9.3.1 Overview of Regional Inequality Indicator	259
9.3.2 Measurement of Inter-county Inequality within Guangdong 1980 – 1990	260
9.3.3 Measurement of Inter-county Inequality within Guangdong 1991 – 2005	267
9.3.4 Regional Inequality within Guangdong: Subjective Measurement	270
9.4 Regional Price Difference in Guangdong	274
9.5 The Issue of Temporary Migration	278
9.6 Regional Inequality Comparison between Guangdong and China	280
9.6.1 General Summary of Intraregional Inequality in the Provinces within China	280
9.6.2 Regional Inequality Measurement in Selected Chinese Provinces	281
9.7 Measurement of Human Development Inequality within Guangdong	285
9.8 Conclusion	295
CHAPTER 10 MAIN CAUSE ANALYSIS OF REGIONAL INEQUALITY IN GUANGDONG: PART I	298
10.1 Introduction	298
10.2 The Historical Factor and Regional Economic Growth	298
10.2.1 The Historical Legacy and Economic Development in Shaoguan	299
10.3 Foreign Direct Investment and Regional Economic Disparity	311
10.4 The Role of the Post-1978 State and Regional Economic Growth	318
10.5 Conclusion	324

CHAPTER 11 MAIN CAUSE ANALYSIS OF REGIONAL INEQUALITY IN GUANGDONG: PART II	325
11.1 Introduction	325
11.2 Geographical Location and Regional Economic Development	326
11.2.1 Coastal Location and Regional Economic Growth	326
11.2.2 Coastal Locational Factor: Limitations	332
11.2.3 Geographical Location and Large Market Proximity	336
11.2.4 Climate and Regional Economic Growth	341
11.3 Industrial Clusters and Economic Agglomeration	346
11.3.1 Measurement of Industrial Clustering	346
11.3.2 The Case of Dongguan: IT Industrial Cluster	354
11.4 Transportation Improvement and Regional Economic Growth	359
11.5 Empirical Model Analysis	370
11.5.1 Geographical Characteristics Analysis	370
11.5.2 Regression Analysis	381
11.6 Final Conclusion	388
CHAPTER 12 GENERAL CONCLUSION	391
APPENDIX	398
Appendix 1: Report of Commercial Price Surveys in Different Regions within Guangdong (Unit: <i>yuan</i>)	398
Appendix 2: Average Annual Sunshine Hours in Various Regions within Guangdong, 1980 – 2004 (unit: hour)	400
Appendix 3: Average Annual Temperature in Various Regions within Guangdong, 1980 – 2004 (unit: °C)	401
Appendix 4: Average Annual Precipitation in Various Regions within Guangdong, 1980 – 2004 (unit: mm)	402
Appendix 5: Preferential Policies of Foreign Direct Investment Designated by the State: Development Type and Date in China	403
Appendix 6: Original Data Set for Regression Analysis	404
Appendix 7: Administration Boundaries of Guangdong's Municipalities (31 December 2004)	408
Appendix 8: Status of Interviewees and Interview Questions	410
BIBLIOGRAPHY	413

LIST OF TABLES

TABLE 2.1: HUMAN AND ECONOMIC DEVELOPMENT IN SELECTED AFRICAN COUNTRIES WITH LOW HUMAN DEVELOPMENT	40
TABLE 2.2: ANNUAL GROWTH RATE OF GDP PER CAPITA IN SELECTED COUNTRIES 1975-2005 (%)	41
TABLE 2.3: CHANGING TRENDS FOR HUMAN DEVELOPMENT INDEX AND PER CAPITA INCOME IN SELECTED COUNTRIES 1975-2005	44
TABLE 2.4: RANK OF GDP AND HUMAN DEVELOPMENT INDICES IN CHINA'S PROVINCES 2003	46
TABLE 2.5: WORLDWIDE COMPARISON BETWEEN HUMAN DEVELOPMENT AND INCOME PER CAPITA, 2005	53
TABLE 3.1: THE AVERAGE CIF/FOB BAND IN SELECTED COUNTRIES 1965-1990 (%)	69
TABLE 3.2: PERCENTAGE OF TOTAL FREIGHT COSTS IN IMPORT VALUE IN AFRICAN REGIONS, 2004 (IN BILLIONS OF DOLLARS)	71
TABLE 3.3: PERCENTAGE OF FREIGHT COSTS IN IMPORT VALUE FOR SELECTED AFRICAN LANDLOCKED COUNTRIES IN THE SELECTED YEARS (IN MILLIONS OF DOLLARS).....	71
TABLE 3.4: THE WORLD RANKING TOP 20 CONTAINER TERMINALS AND THEIR THROUGHPUT, 2003-2005 (UNIT: TEUS MILLION AND THE CHANGE OF PERCENTAGE).....	76
TABLE 4.1: BENEFITS AND COSTS OF INDUSTRIAL CLUSTERS	115
TABLE 5.1: SOURCES OF ECONOMIC GROWTH IN AN ESTIMATION MODEL WITHOUT ADDRESSING HUMAN CAPITAL IN SELECTED REGIONS AND COUNTRIES 1948-1990 ..	157
TABLE 5.2: SOURCES OF ECONOMIC GROWTH IN CHINA, INDIA, AND EAST ASIA DURING THE PERIOD 1978-2004 (ANNUAL PERCENTAGE RATE OF CHANGE)	158
TABLE 6.1: FOREIGN AND DOMESTIC * TRADE IN LEADING SEA PORTS WITHIN CHINA, 1870-1920	176
TABLE 7.1: CRITERIA FOR CONDUCTING INTERVIEW RESEARCH.....	188
TABLE 7.2: ILLITERATE POPULATION AGED 15 AND ABOVE BY SEX AND REGION WITHIN CHINA (2005) (UNIT: PERSON)	190
TABLE 7.3: ILLITERATE POPULATION AGED 15 AND ABOVE BY SEX AND MUNICIPALITIES WITHIN GUANGDONG (2000) (UNIT: PERCENTAGE).....	191
TABLE 7.4: COMPARISON AVERAGE PER CAPITA GDP BETWEEN THREE TARGETED COUNTIES WITHIN GUANGDONG (UNIT: YUAN)	195
TABLE 7.5: AVERAGE PER CAPITA NET RURAL INCOME IN SELECTED PROVINCES OR MUNICIPALITIES IN CHINA 1990, 1995, 2000, AND 2003, (UNIT: YUAN)	197
TABLE 8.1: CHANGE OF TOTAL AMOUNT OF CHINESE GDP AFTER THE 2004 ECONOMIC CENSUS (UNIT: BILLION YUAN)	215
TABLE 8.2: CHANGE OF GDP GROWTH RATE IN CHINA AFTER THE 2004 ECONOMIC CENSUS (UNIT: PERCENTAGE)	215
TABLE 9.1: PER CAPITA NATIONAL INCOME IN DIFFERENT REGIONS WITHIN GUANGDONG 1980-1990, (UNIT: YUAN)	233
TABLE 9.2: PER CAPITA GDP IN VARIOUS REGIONS WITHIN GUANGDONG 1991-2005, (UNIT: YUAN).....	237
TABLE 9.3: PER CAPITA NATIONAL INCOME IN COUNTIES WITHIN GUANGDONG 1980-1990, (UNIT: YUAN).....	242
TABLE 9.4: PER CAPITA GDP IN COUNTIES WITHIN GUANGDONG 1991-2005, (UNIT: YUAN)	248
TABLE 9.5: GINI COEFFICIENT FOR PER CAPITA NATIONAL INCOME IN GUANGDONG'S COUNTIES, 1980-1990 (SHENZHEN INCLUDED).....	261
TABLE 9.6: CV FOR PER CAPITA NATIONAL INCOME IN GUANGDONG'S COUNTIES, 1980-1990 (SHENZHEN INCLUDED).....	262
TABLE 9.7: TOTAL POPULATION IN THE SPECIAL ECONOMIC ZONES OF CHINA 1980-1989 (UNIT: 10 THOUSAND).....	264
TABLE 9.8: CV AND GINI COEFFICIENT FOR PER CAPITA NATIONAL INCOME IN GUANGDONG'S COUNTIES, 1980-1990 (SHENZHEN EXCLUDED)	266
TABLE 9.9: GINI COEFFICIENTS FOR PER CAPITA GDP IN GUANGDONG'S COUNTIES, 1991-2005	267
TABLE 9.10: CV AND GINI COEFFICIENTS FOR PER CAPITA GDP IN GUANGDONG'S COUNTIES, 1991-2005	269
TABLE 9.11: POPULATION DATA VARIATION BETWEEN CENTRAL AND PROVINCIAL	

STATISTICAL DEPARTMENTS IN SELECTED COUNTIES WITHIN GUANGDONG (UNIT: 10,000)	279
TABLE 9.12: HUMAN DEVELOPMENT INDICATORS IN GUANGDONG'S COUNTIES 1981-2000	288
TABLE 9.13: TOTAL GDP IN VARIOUS REGIONS WITHIN GUANGDONG 1995-2004, (UNIT: BILLION YUAN)	296
TABLE 10.1: ANNUAL GROWTH RATE OF GDP PER CAPITA IN SHAOGUAN 1992-2005 (UNIT: %)	299
TABLE 10.2: STATE-ORIENTED CAPITAL INVESTMENT IN LARGE AND MEDIUM DEVELOPMENT PROJECTS WITHIN GUANGDONG, 1957-1964	301
TABLE 10.3: AVERAGE PERCENTAGE OF HEAVY AND LIGHT INDUSTRIES IN TOTAL INVESTMENT OF GUANGDONG'S INDUSTRIES DURING THE PERIOD 1950-1983	303
TABLE 10.4: ROAD TRANSPORTATION DEVELOPMENT IN SHAOGUAN 1949-2006	308
TABLE 10.5: COMPARISON OF THE INDUSTRIAL STRUCTURE BETWEEN SHAOGUAN AND GUANGDONG 2005 (BILLION YUAN)	309
TABLE 10.6: VALUE ADDED TO GUANGDONG'S INDUSTRY, FOREIGN DIRECT INVESTMENT AND EXPORTS 2005	314
TABLE 10.7: PRDA'S SHARE OF GUANGDONG'S EXPORTS, 1980-2004 (US\$ BILLION)	314
TABLE 10.8: TOTAL NUMBERS OF FACTORIES IN GUANGDONG AND THEIR MAINLAND EMPLOYEES, ESTABLISHED BY HONG KONG FIRMS, 2003	316
TABLE 11.1: TRANSPORTATION COSTS FOR HEAVY VEHICLES WEIGHING 10 TONS (YUAN) (CHAOZHOU: A LESS MOUNTAINOUS PERIPHERAL REGION)	331
TABLE 11.2: TRANSPORTATION COSTS FOR HEAVY VEHICLES WEIGHING 10 TONS (YUAN) (XINGNING: A MOUNTAINOUS REGION)	331
TABLE 11.3: TRANSPORTATION COSTS FOR HEAVY VEHICLES WEIGHING 10 TONS (YUAN) (DONGGUAN: A REGION LOCATED WITHIN PRDA)	331
TABLE 11.4: VOLUME OF FREIGHT HANDLED IN MAJOR COASTAL PORTS OF GUANGDONG (UNIT: 10,000 TONS)	335
TABLE 11.5: ROAD DISTANCE BETWEEN DIFFERENT REGIONS IN GUANGDONG (UNIT: KILOMETER)	337
TABLE 11.6: STRUCTURE OF GDP IN GUANGDONG, 1978-2005	344
TABLE 11.7: COMPARISON OF SHARE OF SECONDARY INDUSTRY IN TOTAL GDP BETWEEN REGIONS WITHIN GUANGDONG, 1992-2005 (UNIT: %)	345
TABLE 11.8: VALUE-ADDED INDUSTRIAL OUTPUT OF NINE KEY INDUSTRIES IN GUANGDONG, 2005 (BILLION YUAN)	347
TABLE 11.9: MAIN INDUSTRIES IN THE SPECIALIZED REGIONS WITHIN GUANGDONG	353
TABLE 11.10: ANNUAL GROWTH RATE OF GDP PER CAPITA IN DONGGUAN 1992-2005, (UNIT: %)	355
TABLE 11.11: TOTAL AMOUNT OF EXPORTS AND FDI ACTUALLY UTILIZED IN DONGGUAN 1991-2005, (UNIT: US\$ BILLION)	355
TABLE 11.12: BASIC TRANSPORTATION CONDITIONS IN THE MUNICIPALITIES WITHIN GUANGDONG BY 2004 (UNIT: KILOMETER)	362
TABLE 11.13: ANNUAL GROWTH RATE OF GDP PER CAPITA IN GAOYAO CITY 1992-2005, (UNIT: %)	367
TABLE 11.14: ANNUAL GROWTH OF GROSS INDUSTRIAL OUTPUT VALUE IN GAOYAO CITY 1992-2005, (UNIT: %)	367
TABLE 11.15: SHARE OF FIXED CAPITAL INVESTMENT IN TOTAL GDP IN GAOYAO CITY 1990-2005, (UNIT: %)	368
TABLE 11.16: GEOGRAPHIC CHARACTERISTICS OF THE TWENTY WEALTHIEST COUNTY-LEVEL REGIONS IN GUANGDONG	372
TABLE 11.17: GEOGRAPHIC CHARACTERISTICS OF THE TWENTY POOREST COUNTY-LEVEL REGIONS IN GUANGDONG	374
TABLE 11.18: DESCRIPTIVE STATISTICS	386
TABLE 11.19: COEFFICIENTS (A)	387

LIST OF FIGURES

FIG 5.1: AVERAGE GROWTH RATE OF WORLD PER CAPITA GNP, 1965-90	145
FIG 5.2: AVERAGE PERCENTAGE OF SAVINGS IN GDP	146
FIG 5.3: AVERAGE PERCENTAGE OF INVESTMENT IN GDP	146
FIG 7.1: 2005 RANK OF PER CAPITA GDP IN REGIONS WITHIN MEIZHOU (UNIT: YUAN)	196
FIG 8.1: COMPARISON OF TOTAL GDP OF CHINA BETWEEN PRE-REVISION AND REVISION FIGURES IN 2004 ECONOMIC CENSUS (UNIT: BILLION YUAN)	216
FIG 9.1: PER CAPITA NI IN DIFFERENT REGIONS WITHIN GUANGDONG 1980-1990, (UNIT: YUAN)	234
FIG 9.2: COEFFICIENT OF VARIATION FOR PER CAPITA NATIONAL INCOME IN THE VARIOUS REGIONS WITHIN GUANGDONG 1980-1990 (SHENZHEN EXCLUDED)	235
FIG 9.3: PER CAPITA GDP IN VARIOUS REGIONS WITHIN GUANGDONG 1991-2005, (UNIT: YUAN)	238
FIG 9.4: COEFFICIENT OF VARIATION FOR PER CAPITA GDP IN VARIOUS REGIONS WITHIN GUANGDONG 1991-2005	239
FIG 9.5: GINI COEFFICIENT AND CV FOR PER CAPITA NATIONAL INCOME IN GUANGDONG'S COUNTIES, 1980-1990 (SHENZHEN INCLUDED)	262
FIG 9.6: CV AND GINI COEFFICIENT FOR PER CAPITA NATIONAL INCOME IN GUANGDONG'S COUNTIES, 1980-1990 (SHENZHEN EXCLUDED)	266
FIG 9.7: CV AND GINI COEFFICIENTS FOR PER CAPITA GDP IN GUANGDONG'S COUNTIES, 1991-2005	269
FIG 9.8: SURVEY 1	273
FIG 9.9: SURVEY 2	273
FIG 9.10: SURVEY 1	273
FIG 9.11: SURVEY 2	274
FIG 9.12: COMPARISON OF THE GINI COEFFICIENT FOR PER CAPITA INCOME FOR COUNTIES IN GUANGDONG, SHANDONG, HENAN AND GUIZHOU IN 1988, 1992, 1999 AND 2004	283
FIG 9.13: COMPARISON OF THE CV FOR PER CAPITA NI AND GDP BETWEEN GUANGDONG, SHANDONG, HENAN AND GUIZHOU IN 1988, 1992, 1999 AND 2004	284
FIG 9.14: RATIO OF PEOPLE AGED 65 AND ABOVE TO TOTAL POPULATION IN GUANGDONG, SHANDONG, HENAN AND GUIZHOU 1980-2000, (UNIT: %)	293
FIG 9.15: CRUDE DEATH AND INFANT MORTALITY RATES IN COUNTIES WITHIN GUANGDONG, SHANDONG, HENAN, GUIZHOU AND NATIONALLY, 2000 (UNIT: ‰)	294
FIG 9.16: ANNUAL GROWTH RATE OF PER CAPITA GDP IN GUANGDONG 1979-2005, (UNIT: YUAN)	295
FIG 10.1: ANNUAL INDUSTRIAL GROWTH IN SHAOGUAN, 1982-2005 (%)	309
FIG 10.2: TOTAL AMOUNT OF UTILIZED FOREIGN DIRECT INVESTMENT IN GUANGDONG, 1979-2005 (US\$ BILLION)	313
FIG 10.3: MAIN TRADE FORMS OF IMPORTS AND EXPORTS BY THE FOES LOCATED IN GUANGDONG 2005, (UNIT: PERCENTAGE)	318
FIG 10.4: THE FIVE MAIN REASONS GIVEN BY HONG KONG COMPANIES FOR OPENING FACTORIES IN GUANGDONG	322
FIG 11.1: FIELD SURVEY 3 (UNIT: PERCENTAGE)	328
FIG 11.2: SHARE OF THE INDUSTRIAL OUTPUT VALUE OF ELECTRONIC INFORMATION INDUSTRY IN GUANGDONG BY THE TOP THREE REGIONS, 2005	348
FIG 11.3: SHARE OF THE INDUSTRIAL OUTPUT VALUE OF ELECTRONIC MACHINERY AND SPECIAL PURPOSE EQUIPMENT INDUSTRY IN GUANGDONG BY THE TOP THREE REGIONS, 2005	348
FIG 11.4: SHARE OF TOTAL INDUSTRIAL OUTPUT VALUE OF PETROLEUM AND CHEMICAL INDUSTRY IN GUANGDONG BY THE TOP THREE REGIONS, 2005	348
FIG 11.5: SHARE OF TOTAL INDUSTRIAL OUTPUT VALUE OF TEXTILE AND GARMENTS INDUSTRY IN GUANGDONG BY THE TOP THREE REGIONS, 2005	349
FIG 11.6: SHARE OF TOTAL INDUSTRIAL OUTPUT VALUE OF FOOD AND BEVERAGES INDUSTRY IN GUANGDONG BY THE TOP THREE REGIONS, 2005	349
FIG 11.7: SHARE OF TOTAL INDUSTRIAL OUTPUT VALUE OF BUILDING MATERIALS INDUSTRY IN GUANGDONG BY THE TOP THREE REGIONS, 2005	349

<i>FIG 11.8: SHARE OF TOTAL INDUSTRIAL OUTPUT VALUE OF TRANSPORT EQUIPMENT MANUFACTURING INDUSTRY IN GUANGDONG BY THE TOP THREE REGIONS, 2005</i>	351
<i>FIG 11.9: SHARE OF TOTAL INDUSTRIAL OUTPUT VALUE OF LOGGING AND PAPERMAKING INDUSTRY IN GUANGDONG BY THE TOP THREE REGIONS, 2005</i>	351
<i>FIG 11.10: SHARE OF TOTAL INDUSTRIAL OUTPUT VALUE OF MEDICINE INDUSTRY IN GUANGDONG BY THE TOP THREE REGIONS, 2005</i>	351

LIST OF MAPS

<i>MAP 9.1: THE FIVE COUNTIES IN GUANGDONG WITH HIGHEST PER CAPITA NI IN 1990</i>	241
<i>MAP 9.2: DISTRIBUTION OF PER CAPITA NATIONAL INCOME IN COUNTIES WITHIN GUANGDONG, 1980 (UNIT: YUAN)</i>	255
<i>MAP 9.3: DISTRIBUTION OF PER CAPITA NATIONAL INCOME IN COUNTIES WITHIN GUANGDONG, 1990 (UNIT: YUAN)</i>	256
<i>MAP 9.4: DISTRIBUTION OF PER CAPITA GDP IN COUNTIES WITHIN GUANGDONG, 2000 (UNIT: YUAN)</i>	257
<i>MAP 9.5: DISTRIBUTION OF PER CAPITA GDP IN COUNTIES WITHIN GUANGDONG, 2005 (UNIT: YUAN)</i>	258
<i>MAP 11.1: SPATIAL DISTRIBUTION OF THE TWENTY RICHEST REGIONS IN GUANGDONG, 2005</i>	378
<i>MAP 11.2: SPATIAL DISTRIBUTION OF THE TWENTY POOREST REGIONS IN GUANGDONG, 2005</i>	379

Abbreviations

FDI	foreign direct investment
NI	national income
GDP	gross domestic products
TFP	total factor productivity
HDI	human development index
PRDA	Pearl river delta area
NEG	new economic geography
CV	coefficient of variation
NGO	non-governmental organization
NSB	national statistical bureau of China
GSB	Guangdong statistical bureau

Chapter 1 General Introduction

1.1 The Issue of Regional Inequality

Do regional inequality and unbalanced regional development really matter? The report published by the United Nations Development Programme (UNDP) (2000) suggests the crucial role played by equality in ensuring the development of human well-being and national welfare. However, the world has been facing the challenge of inequality in spatial and human development. Numerous countries suffer from the problems brought about by rising regional inequality, such as social unrest, poverty and high crime rates and for some nations these outcomes are extremely severe. The full spectrum of inequality lies in the comprehensive and multidimensional framework of development itself. Turner (1986) and Arcelus, Sharma and Srinivasan (2006) have offered detailed discussion regarding multidimensional development. There are various aspects of inequality, such as income, education, health, human development, and political participation. Human diversity has a natural impact on these aspects and the study conducted by Sen (1992) points out the significance of advantages and disadvantages inherited by human beings; therefore, it is necessary to distinguish the types of regional inequality mainly caused by human diversity.

Sen's argument is important. The concept and boundaries of inequality are difficult to define with accuracy. When dealing with the issue of regional inequality, the type of inequality under discussion should be clearly highlighted. As far as this study is concerned, regional income disparity is the key issue I am keen to explore and analyze. Fields (1989) has provided a valid definition of income inequality.

Income and wealth are the fundamental benchmark indicators for measuring inequality. This argument is based on two reasons. First, income growth is a good measure of economic development. Secondly, income has been the determining influence on human development: income growth establishes the key foundation for development of human well-being and related social aspects. Nevertheless, other significant aspects of disparity in human development will be addressed in this research, such as education, health and life expectancy.

McGillivray's study (2006) demonstrates the unequal development of public health and other social aspects between different countries and highlights the core discussion issues of regional diversity and human development inequality within the international community. Numerous scholars have completed significant studies of the serious effects of rising regional inequality faced by many nations. On the one hand, worsening regional inequality badly affects long-term economic growth and impacts negatively on economic development (e.g. Pedroni and Yao, 2006). On the other hand, more significantly, rising inequality threatens social and political stability and national utility. As Fields (2001, p. 201) states, "High inequality contributes to political and macroeconomic instability." Muller (1985) states that policymakers should consider the issue of regional equality when they implement strategies of national economic development: otherwise, severe political violence will inevitably occur. Drèze and Sen's study (1995) has also raised this issue, arguing that uncontrolled regional inequality would not only affect human development potential, but also more significantly, cause political instability, social conflict and violence.

Their study cites the cases of the severe regional disparity within Brazil, other areas of America and South Africa.

The issue of spatial inequality has increasingly attracted attention from academics. According to the World Development Report (1996, p. 70), for example, in June 1995, 20 percent of the Russian territories (the resource rich regions and Moscow) produced over 40 percent of national total income. However, the poorest 20 percent of territories (e.g. ethnic republics) produced only 5 percent of total income in Russia. Ray's study (1998) found that income per head in the richest nations (Switzerland) was, strikingly, nearly 400 times greater than in the poorest (Tanzania). In an early study, Williamson (1991) pointed out that America had also suffered from the problem of rising regional inequality during the rapid economic development of the nineteenth century. Some Latin American studies focus on worsening regional income disparity within Brazil (e.g. Fields, 2001). Brazilian national wealth has been concentrated in the southeastern region surrounding São Paulo where average income per head is much higher than in other regions. This regional income disparity has faced Brazilian society with challenges, such as high crime rate, social conflict, and poverty. As far as Asia is concerned, Nepal is a classic example of severe regional inequality leading to political instability and bloody civil war. The serious Maoist insurgency, social unrest and violent conflict in Nepal were originally caused by rising social and economic disparities among the different regions and have negatively affected the country's national industrial and economic development (Asian Development Bank 2007, 2007). The study conducted by Murshed and Gates (2005) suggests that regional inequality between capital Kathmandu and western districts is particularly significant. For example, the average income per capita in Rolpa, Achham, and Salyan accounted for only 25, 24 and 17 percent respectively of the equivalent figure in Kathmandu in 1996.

As far as China is concerned, during the reform period, rising regional income inequality has been widely discussed (e.g. Chen and Fleisher, 1996; Yao and Zhang, 2001; Renard, 2002). Severe spatial disparity has created a major challenge to the Chinese government's attempts to maintain balanced economic development, social justice and political stability. For example, Chen and Fleisher (1996) argue that rising regional income inequality would not only cause social unrest and political instability, but also would disrupt improvements to China's long-term prospects. Furthermore, research by Yao and Zhang (2001) suggests unequal regional development and rising income disparities between the developed coastal, moderately-developed central and least-developed western regions within China. They further stress that the potentially negative impact of a failure to address such regional disparity on China's future prosperity.

Wang and Hu (1999) conduct an illuminating analysis of the potential consequences of regional economic disparity in China. Indeed, theirs is probably the strongest and most straightforward such study since 1978. They stress that spatial inequality is a matter of political importance; and the historical relationship between regional economic disparity, social stability and national political unity is very close. Wang and Hu (1999) claim that a catastrophic political outcome is the natural consequence of severe spatial income inequality. They state that:

China's own history is full of uprisings, rebellions, and revolutions

sparked by economic injustice. Even if the crisis does not reach the point of revolution, social tension and instability are harmful to economic growth. Thus, there is a great danger that growing inequality will derail China's reforms and imperil its future growth. (1999, p. 201)

In their discussion of China, Wang and Hu (1999) further add that it is unjustified and difficult to accept the fact of rising spatial income inequality in terms of the moral legitimacy of a country established on fundamental egalitarian principles. Furthermore, Pedroni and Yao (2006) have pointed out the serious challenge for the Chinese government posed by rising regional inequality.

According to Kuznets (1953), achieving fast economic growth simultaneously in all countries is not possible. Due to the difficulties in adopting and imitating advanced production methods and new economic mechanisms facing the under-developed nations, regional income disparity between the economically rich and poor countries is the likely outcome. Therefore, economic development is necessarily linked to unbalanced regional growth and worsening income inequality among different nations. This is the U-curve hypothesis of regional disparity proposed by Kuznets. Kuznets's thought will be discussed in detail in the following chapter. Nevertheless, the evidence presented earlier from selected countries suggests that income inequality makes a difference to economic growth and human development and must be taken seriously, if long-term economic growth and sustainable human development are to be achieved. As inequality can also negatively affect social harmony and political stability within a country, its causes and consequences must be effectively addressed and resolved. Rising inequality is having an epidemic effect on long-term development and social stability, which, uncontrolled, would affect social welfare and fulfillment of human potential. In terms of China, the issue of regional development disparity is very prominent and has presented an increasing challenge during the implementation of 'reform and open-door' policies over the past two decades. The extent of regional inequality has reached a striking level, and has been recorded as one of the highest regional disparities in Asia: second only to Nepal (Asian Development Bank 2007, 2007).

Regional economic disparity can not only lead to political instability and social conflict, but also, more significantly, can trigger the severe consequence of human health problems. Botswana being a good example of this issue. Although Botswana has made impressive economic achievements over the last few decades, internal regional income inequality has simultaneously been widening (Clover, 2003): accompanied by deteriorating health conditions. The average life expectancy at birth had decreased to only 44.7 years in 2001 from 68 years in 1966 due to the prevalence of the HIV/AIDS epidemic (Clover, 2003) from which many people living in Botswana have suffered. An important report '*AIDS Epidemic Update: Special Report on HIV Prevention*', UNAIDS and WHO (2005) suggests the severe situation relating to the HIV/AIDS epidemic in this country: up to six percent of the children aged between 18 months and four years old were infected by HIV; 29 and 21 percent of adults aged 45-49 years, and in their early 50s respectively, were found to be HIV-positive. In addition, the '*2005 Botswana Second Generation HIV/AIDS Surveillance: Technical Report*', Ministry of Health, Republic of Botswana (2005, p.18) admits that:

HIV/AIDS in Botswana has become a serious developmental and social problem impacting negatively on all sectors.

Although regional inequality was already high in 1978, during the reform period¹, China has been facing widening economic inequality between the eastern, central and western regions. In contrast to the economically prosperous and developed eastern region, the western region has recorded much lower economic growth and remained in a relatively backward position, with the slow economic development and low incomes leading to poor standards of human health. For example, the average life expectancy in the less developed western region is 68.4 years in contrast to 71.9 years in the developed eastern region. In certain under-developed and mountainous regions in the west, there are no local clinics and hospitals (UNDP, 2005). In addition, the UNDP (2005, p. 58-59) suggests that regional health inequality can be revealed by the total number of hospital beds. For example, the average number of hospital beds per thousand people is less than two in poor Guizhou in contrast to the equivalent figure of six in wealthy Shanghai. This report highlights the much faster improvement in medical provision in the eastern area compared to central and western areas during the reform period. In 2002, in the eastern region, the number of medical personnel per 1,000 residents was 3.8; however, the equivalent figures were just 3.2 and 3.0 in the central and western regions respectively. In terms of the number of assistant doctors per 1,000 residents, in contrast to 1.4 in both the central and western regions, the equivalent figure was 1.7 in the east (UNDP, 2005, p. 59). These figures illustrate that regional health differences and population health inequalities are widespread in China.

Some scholars argue that China has been a victim of the co-existent challenges of rapid economic growth and slow human development. Haq (1999) cites the case of China in a striking comparison between fast economic growth and deterioration in human conditions. Sanjay Reddy (2007, p. 50) highlights precisely the same point:

In the light of China's advances in poverty reduction, it is important to assess whether it has made commensurate improvements in other spheres of human well-being, such as education, health and access to basic services.

The evidence from Botswana and China suggests that there is close correlation between widening regional income disparity and worsening population health. Based on his research on America, Richard Wilkinson (1999) has examined the relationship between regional income inequality and population health. He stresses that a widening regional income gap is closely linked with social conflict and worsening population health: an extremely hierarchical income structure will lead to poor health conditions. In his early study, Wilkinson (1996) points out that psychosocial factors directly influence population health. Meanwhile significantly, human psychosocial stress and difficulties are mainly caused by low income and poverty. Consequently, he claims that psychosocial factors play an important role in illustrating the significant influence of wealth and income growth on people's health.

¹ In terms of my research, the reform period which I specifically discuss, is the years between 1978 and 2005.

In general, reduction of regional income inequality would improve human health conditions for all groups in society: both rich and poor. In addition, by narrowing income disparities, the scale of poverty among the poor would be reduced and their relatively bad health conditions could be improved: both of which would contribute to increasing a nation's average life expectancy (Wilkinson, 1996). Moreover, a study conducted by Kennelly, O'Shea, and Garvey (2003) maintains that wealth and income do make a difference to a population's health in that income per capita has considerable influence on human health outcomes. They maintain that income per capita has big influence on shaping human health outcomes. In his research, Wilkinson (1990; 1996) has cited evidence from the developed countries to demonstrate the positive relationship between good national health (shown in life expectancy) and an egalitarian society. Conversely he argues through an empirical comparison of different countries, that unequal regional income distribution leads to high national mortality rates. He quotes the examples of Japan, which since World War II, has achieved low income inequality, national health improvement (e.g. the rise of life expectancy) and rapid overall social development, and Britain, which during the period between World War I and II, achieved health improvements due to social cohesion and regional income equality (Wilkinson, 1996).

However, there is another argument relating to the issue of regional disparity which should not be ignored. Some scholars believe that, state efforts to promote regional economic equality might impact negatively on overall economic growth and that regional disparity of opportunity is an unresolvable issue. For example, Alonso (1968) argues that a government's efforts to achieve regional equality in economic development can result in a slow-growing economy without in fact lessening regional disparity because these policies embody both efficiency and equity goals², which can be contradictory. According to neoclassical economic theory, regional income disparity will necessarily exist, and may provide economic opportunities during the early development stage when enhancement of private incentives and improvements in efficiency are particularly desirable. Regional income equality might, indeed, cause slower improvement of overall human welfare and various scholars (e.g. Stiglitz, 1969; Bourguignon, 1981; Mirrlees, 1971; Aghion, Caroli, and García-Peñalosa, 1999) have provided detailed discussion regarding this issue.

Nevertheless, the experience of many nations reveals that regional disparity is detrimental to economic growth and human welfare development. Nations exhibiting internal inequality have slower economic development than countries with less regional disparity. For example, relative to the Philippines, South Korea has recorded relatively low income disparity among the different population groups since the early 1960s. Significantly, economic development in South Korea was much faster than in the Philippines between the 1960s and 1990s. (Aghion, Caroli, and García-Peñalosa, 1999)

My view is as follows: regional inequality makes a difference to economic growth and widening spatial income disparities should be resolved. Severe regional inequality will negatively influence social harmony and national stability, and more significantly, overall human development. As the evidence discussed earlier demonstrates, regional

² "By efficiency is meant, most simply, national economic growth, often measured in terms of per capita national product or a discounted consumption stream. By equity is meant a more equal distribution of income." (Alonso, 1968, p. 1)

income equality plays a crucial role in ensuring political stability and population health improvement. As far as China is concerned, and as will be discussed in detail in the following chapter, the issue of worsening regional economic disparity during the reform period is striking and requires urgent resolution by the state. Both the rate of economic growth and long-term development will be negatively affected by spatial income inequality. Moreover, the fundamental egalitarian principles and moral justice of this so-called socialist country would be severely damaged if spatial disparity were not to be addressed. It is hard to maintain social cohesion and harmony when regional income inequality and disparities in human development are widening.

However, it is not my intention to suggest that the current regional inequality will inevitably cause social revolution and civil war in China. In fact, it is far from it. The severity of spatial inequality has not reached the point of triggering social revolution and bloodshed. Nevertheless, personally, I do think that widening regional disparity is a real threat to China and a serious challenge to the maintenance of long-term economic growth. More significantly, without dealing with regional inequality, many other problems, such as poverty, and underdevelopment of health and education, will remain unresolved. If the state attempts to ignore the regional inequality problem, this could give rise to social revolution and rural rebellion in the future. This is not pure speculation or subjective prediction as the social unrest in Xinjiang and Xizang illustrates. In addition, some scholars believe that during China's five-thousand-year history, many revolutions and rural rebellions were directly related to, or, caused by, severe income disparity.

1.2 The Significance of Guangdong

This study attempts to systematically analyze the case of Guangdong Province to reflect and overview the issue of regional inequality in China during the post-1978 period. Guangdong, which is located in the southern region and adjacent to the coast, is one of the key industrial centres and economic powerhouses in China and has been enjoying impressive economic development since late 1978, taking the lead in many aspects, such as the foreign direct investment, total GDP, and foreign exports and imports. This province has, indeed, been a reform pioneer and laboratory region in terms of economic development. Studies by the Chinese scholars, Li (2006) and Mei (2006) describe the rapid growth of foreign trade, foreign direct investment and GDP in Guangdong during the reform period, especially during the last ten years. Moreover, the studies conducted by Vogel (1989) and Sung, et al, (1995) provide, detailed discussion of the advanced and economically powerful position of Guangdong in China. According to official statistics, some of its key outstanding performance in economic development can be demonstrated as follows:

- The total value of foreign exports increased to 238.2 US\$ billion in 2005 up tenfold from 22.2 US\$ billion in 1990, and in 2005, Guangdong accounted for 31.2 percent of total value of foreign exports in China
- The GDP in Guangdong jumped to 2236.6 billion *yuan* in 2005, which accounted for more than 12 percent of the total GDP in China (18308.4 billion *yuan*)
- Guangdong has been the key focus of foreign direct investment in China. Actually utilized FDI in Guangdong jumped to 15.2 US\$ billion, multiplying 168 times from 0.09 US\$ billion in 1979, Guangdong's share of China's FDI

was up to 23.8 percent

- Per capita GDP in Guangdong was 24438 *yuan* in 2005, which was more than one and a half times higher than the equivalent figure in China (14040 *yuan*)

Guangdong was one of the first provinces in China to exercise and implement the so-called 'reform and open-door' policies issued by the central government to transform its traditional planning economy to modern market economy and to develop local economies. Various state-oriented preferential policies have been extended to Guangdong during the reform period leading Cheng (1998) to suggest that it is a laboratory and testing region for governmental development policies. Furthermore, Li and Yang (2003) mention the setting up of Special Economic Zones and testing of state-oriented reform policies in administrative and economic sectors. In that sense, Guangdong is one of the key change makers during the reform period. Guangdong is a pioneer province in China, instigating economic reform and opening up to the outside world.

In order to understand reform and opening-up policies, and analyze their impact on the regional economic development in China since 1978, without question, Guangdong is the best site available for my research, which will also reveal the impact of such policies on the regional development of other Chinese regions. Owing to its development characteristics and prominence as a region since 1978, changing tendencies of spatial disparity in Guangdong can generally reflect the situation in many other Chinese regions. This is especially true of the eastern and coastal regions, which have similar development strategies: emphasizing export-oriented industries, foreign direct investment and privately-led economic growth. In other words, Guangdong is a mirror reflecting overall regional disparity. Moreover, the study of Guangdong can provide some useful lessons regarding regional economic growth and implications of spatial economic disparity for other regions in the world, for example, Vietnam in Asia and Egypt in Africa.

To my mind, this is an exciting place for regional economic study. Guangdong is densely-populated with both coastal developed and inland under-developed regions; and the newly established Special Economic Zones (e.g. Shenzhen and Zhuhai) and traditional industrial centre (Guangzhou). Moreover, Guangdong has export- and industry-oriented, rapidly developing regions (e.g. Dongguan) as well as areas mainly engaging in agriculture (e.g. Wuhua and Heping), making it representative of the overall diversity in China. Guangdong is, however, unique in its sharp spatial economic differences between the coastal and mountainous regions. Spatial inequality research in Guangdong will certainly benefit our knowledge and enrich our understanding of unbalanced regional economic growth in the context of China. In addition, the regional development phenomenon of Guangdong will allow us to explain and test theories and models of regional economic growth, divergence and convergence.

In fact, the pioneer status of Guangdong has existed for a long time. Numerous early revolutionaries and reformers in China, such as Kang Youwei and Liang Qichao, are native Cantonese. Sun Zhongshan, the key founder and first President of the Republic of China, was a distinguished Cantonese native. Sun was also the main leader of the 1911 Revolution which directly caused the end of the Qing dynasty. Moreover, the

significance of Guangdong in contemporary Chinese history is evident. Mao Zedong, who was the founder of the People's Republic of China, won key support from the peasants and farmers in the rural regions. However, Mao made his first move and initially promoted his revolutionary policies from the Peasant Training Institute in Guangdong. The importance to China of Huangpu Military Academy established in Guangdong, is also widely recognized. Both Vogel (1989) and Sung, et al, (1995) have carried out important research on the significant status of Guangdong in Chinese history.

Moreover, although numerous previous studies regarding this province have been conducted (e.g. the city study of Dongguan conducted by Yeung, 2001), the issue of regional economic development and the causes of spatial disparity in Guangdong have yet to be systematically addressed. This is especially a particular need to update measurements of regional disparity and main cause analysis. Many scholars have concentrated on research into the so-called Guangdong development model, mainly based on state development policies, the large amount of FDI, and discussion of economic success of Guangdong (e.g. Sung, et al, 1995; Cheng, 1998; Wu, 1999; Bui, et al, 2003).

There are previous studies of spatial disparity within Guangdong Province (e.g. Ouyang, 1993; Fan, 1995, Weng, 1998; Gu, et al, 2001); and Ouyang's (1993, cited in Gu, et al, 2001) is probably one of the earliest of these. Ouyang points out that the inter-regional economic disparity between the Pearl River Delta Area (PRDA) and other regions in Guangdong widened during the 1980s in terms of GDP per capita; however, he argues that intra-regional inequality within PRDA had remained stable. In addition, Fan (1995a; 1995b) found that inter-county levels of economic inequality had intensified during the period from 1982 to 1990, as shown by the rapid development in the PRDA. By contrast, she believes that economic regional disparity decreased between 1990 and 1993 due to the growth of township and village enterprises and increasingly even distribution of FDI among regions within Guangdong.

Weng (1998) argues that the regional economic disparities within the PRDA had actually declined due to the slow growth and weakening power of traditional large cities like Guangzhou and rapid economic development in many backward and under-developed regions within the PRDA such as Doumen and Gaoming. In general, Weng's research suggests that the regional inequality within the PRDA has been narrowed since 1978. My geographical research area is larger than Weng's as he only covers the PRDA, rather than the whole of Guangdong. In contrast, regional disparity between the PRDA and the remainder of Guangdong is emphasized by my study. This research intends to argue that the regional disparities within Guangdong have been widening during the reform period. Nevertheless, I disagree with Weng's argument regarding the economic decline of Guangzhou and detailed discussion on this topic is presented in Chapter Ten.

Wang and Hu (1999) suggest that the problem of spatial inequality within Guangdong is severe, although Shen, Wong and Chu (2001) maintain that regional inequality within Guangdong experienced a fluctuating upward trend during the period from 1980 to 1998. The CV for GDP per capita reached 1.14 in 1985 from 0.78 in 1980; although the CV figure was down to 0.87 in 1998, it was much higher than the

equivalent figure in 1980 (Shen, Wong and Chu, 2001). The study conducted by Gu, et al, (2001) is probably the most detailed and recent research into the issue of regional inequality within Guangdong during the reform period. By adopting the various social and economic indicators, their study suggests that Guangdong has been facing uneven regional development during the reform period; the pace of economic growth in the PRDA has been particularly fast. Gu, et al, (2001, p. 110) state that:

For the province as a whole, the concentration of economic power in the Zhujiang Delta region means a spatial polarization process. At the same time, the relative economic strength of peripheral areas began to decline.

Relative to some other regions in China, such as Zhejiang and Jiangsu, less attention has been given to Guangdong in terms of intra-regional economic study and cause analysis. I attempt to conduct detailed studies on this province in order to bridge this research gap and hope that my research can provide, useful and illuminating discussions.

1.3 Research Structure Overview

The aims and objectives of this research are as follows. First, I attempt to provide a detailed and updated measurement of regional inequalities in Guangdong during the reform period, including both economic and human development aspects. Secondly, more significantly, I hope to offer a proper estimation of the significance of the effect of geography on regional economic development in Guangdong. This research attempts to provide an appropriate conclusion regarding the key causes of unbalanced regional economic development in this southern province. What contribution has the geographical factor made to development of the core – periphery pattern of spatial disparity in Guangdong? Has geography inevitably led to the developmental tendency of worsening regional inequality since 1978?

In order to achieve my research aims, two main hypotheses are proposed by this study. Hypothesis one suggests that regional economic inequality within Guangdong has widened during the reform period. Hypothesis two proposes that the worsening regional inequality within Guangdong is mainly caused by its geography. The methodological approach adopted by this study will be discussed in detail in Chapter Seven but I intend to address and test the two hypotheses, using as my main approaches, field interviews and official data analysis.

As previously outlined, there are three main parts to this PhD thesis. The first part discusses the theoretical literature and relevant frameworks for measurement of regional inequality, traditional economic geography, economic convergence and divergence theories are presented. A general literature survey regarding both Chinese and worldwide evidence is also included. The second part of this thesis emphasizes measurement of regional income inequality within Guangdong during the reform period, including measurement of human development disparities within the different regions. The third part analyzes the main causes of widening regional inequality within Guangdong. The detailed structure of this thesis is as follows:

Chapter Two will provide general literature discussion regarding measurement of inequality. The two most popular approaches, the Human Development Index (HDI)

and Gross Domestic Products (GDP), will be analyzed in this chapter as well the Happiness theory for measuring human development. The key issues addressed by this chapter are as follows: Is it really justified to argue that the GDP approach has been useless in the analysis of regional inequality? Relative to the income approach, is the HDI a better approach to measurement of overall regional inequality and human development? How useful is the Happiness thought measure of disparity in human development? Chapter Three offers the literature framework, part II on traditional geographical theories and their importance in assessing regional economic development. I will also investigate the following important issue: what is the role played by geographical location and market proximity in unbalanced regional development and regional income disparity?

Chapter Four will provide analysis of the theoretical literature, part III on regional economic divergence. Numerous key academic theories on regional divergence, including those of Myrdal, Friedmann and Krugman, are discussed in this chapter. What is the impact of economic agglomeration on unbalanced industrial distribution among the regions? How significantly does industrial cluster affect the core – periphery pattern of regional economic disparity? These are the key issues which are addressed in this chapter. Chapter Five presents the literature on regional convergence theories. Kuznets's inverted U-curve theory of regional economic development and income disparity and neoclassical growth theories are discussed. Moreover, worldwide evidence concerning the development tendency of regional income disparity is assessed in this chapter.

Chapter Six presents a general literature survey of the effects of geographical factors on regional economic development in China and the general change in national trend of spatial income inequality. The macroregional model developed by Skinner is also analyzed in this chapter. The main issues are these: what is the contribution made by location and market proximity to regional industrial and economic development? What are the impacts of industrial cluster and agglomeration economies on shaping regional development? What is the developmental tendency of regional income disparities within China?

Chapter Seven presents fieldwork research design and research methodology adopted in this study. The research methods, sample, reliability and ethical issues are discussed in this chapter. Chapter Eight offers detailed discussion on the quality and reliability of Chinese statistical data; and the capability of official statistical departments to handle such data. Moreover, the specific statistical quality in Guangdong is addressed.

Chapter Nine focuses on measurement of regional inequality within Guangdong during the reform period. Various economic and human indicators have been adopted to measure this developing trend. This chapter offers a comparable and time-series measurement of regional economic development and spatial inequality within Guangdong since 1980. I have adopted two spatial frameworks for regional inequality analysis which are at inter-regional and inter-county levels. The majority of statistical data was collected from official sources such as the statistical yearbook, policy documents and government working reports. General comparison of the intra-regional inequality between Guangdong and China is presented in this chapter. Three

representative provinces – Guizhou, Henan and Shandong – have been chosen for detailed regional disparity comparison with Guangdong.

The main cause analysis of regional inequality within Guangdong has been separated into two parts. Chapter Ten Part I analyzes the importances of factors boosting regional economic development, such as state, history and foreign direct investment. Chapter Eleven Part II investigates how location, industrial agglomeration and transportation costs can cause regional economic growth and regional inequality in Guangdong. Various representative regions have been chosen for detailed analysis. Empirical model analysis is also demonstrated in Chapter Eleven, including geographical characteristics and regression model analysis. The following specific questions are highlighted. What is the contribution made by the geography to shaping the core – periphery pattern of spatial disparity? How significantly has geography influenced local economic growth since 1978? Has geography been the driving force behind the tendency for worsening regional inequality within Guangdong during the reform period?

There are eight Appendices contained in this research. Appendix 1 provides a report on a commercial price survey of different regions within Guangdong which was conducted by the writer during field research. Appendices 2, 3, and 4 offer specific meteorological data from the various regions within Guangdong, including annual sunshine hours, annual temperature and annual precipitation. Appendix 5 describes special state-oriented FDI policies in various regions within China during the reform period. Appendix 6 shows the original data set for regression analysis. Appendix 7 defines the administrative division of Guangdong. Finally, Appendix 8 provides a list of interviewees' details and interview questions.

Chapter 2 Literature Review Part I: Inequality Measurement Approaches

2.1 Introduction

The first main hypothesis proposed by this research is to measure regional inequality within Guangdong. Therefore, this chapter provides a general review of literature relating to inequality measurement approaches. In particular, two widely used approaches – the Human Development Index (HDI) and Gross Domestic Product (GDP) will be critically discussed. The Happiness theory developed by Layard for human development measurement, will also be specifically addressed. The main questions I attempt to explore are these: In analyzing the issues of regional inequality and human development disparities, is it really justified to argue that the GDP approach is out of date and should be abandoned? Does the HDI offer a real alternative to GDP in measuring overall inequality and human development? In addition, relative to the HDI and GDP approaches, does Happiness thought offer a reliable and new approach to measurement of human development? These important questions will be addressed in this chapter.

The structure of the chapter is as follows. Section Two focuses on discussion of advantages and disadvantages of the GDP and HDI measurement approaches citing evidence from worldwide sources. In addition, the advantages of using a single indicator, such as life expectancy at birth, for measuring human development disparity are highlighted in this section. Section Three attempts to analyze the importance of Happiness Theory to understanding human development. Section Four draws conclusions.

2.2 Inequality Measurement Approaches: GDP and HDI

Rising regional inequality is a serious problem concerning to many nations. What is the best approach to assessing and analyzing regional disparity? Whilst various approaches to inequality measurement have been critically discussed by the international academic community, dramatically differing opinions have been expressed (e.g. Lewis, 1955; Ahluwalia, 1976; Sen, 1999; Haq, 1999; Fields, 2001). Heated academic debate is mainly concentrated on two such approaches – Gross Domestic Product (GDP) and Human Development Index (HDI), with the former attracting particularly strong criticism.

2.2.1 GDP Measurement Approach: Advantages and Criticism

In a very early study, Arthur Lewis (1955) argued that,

Economic growth also gives us freedom to choose greater leisure. In the primitive state we have to work extremely hard merely to keep alive. With economic growth we can choose to have more leisure or more goods ... (p. 421)

Lewis's research is reasonable and meaningful. It stresses the importance of stable economic growth and wealth in material production to ensuring human development. Without economic growth, development of human potential would be extremely difficult to achieve, as economic growth and human development are two sides of the same coin. Human development and freedom must be based on economic growth and plentiful material resources. It is hard to imagine an economically poor person having freedom of choice, or the ability to develop to his/her full potential. Similarly, it is almost impossible for an economically backward region to spend more money on improving human living standards and developing social welfare. Although other factors are involved in international income disparities, Ahluwalia (1976) rightly claims that the importance of income per capita in shaping economic inequality should not be underestimated.

Economic development can be best demonstrated by income growth (e.g. GDP per capita). As Fields (2001, p. 2) highlights, "Economic development is sought primarily via economic growth." The inherent advantage of GDP approach lie not only in its measurement simplicity and reliability, but also in the ease of income data collection and analysis. The GDP approach analyzes the key indicator affecting regional inequality and the daily lives of human beings – total income or income per capita. As Greg Mankiw (2006, p. 17) states,

From either viewpoint, it is clear why GDP is a gauge of economic performance. GDP measures something people care about - their incomes.

Therefore, by measuring income growth and economic development, the real extent of regional income disparity can be properly understood. Moreover, an overall picture of regional development within a country can be generally estimated as income per capita is also probably one of the best single indicators of aspects of social development disparity, such as education, health and environment. The fairness and credibility of the GDP approach is illustrated by the latest study conducted by the Asian Development Bank (2007) which suggests the seriousness of regional development disparity within Nepal by adopting the GDP measurement approach. Average expenditure per capita has increased to 30 percent in the capital city, Kathmandu and eastern Terai region compared to only 5 percent growth in the Eastern Hills Area of Nepal. Furthermore, Brazil is a country which has been facing extremely serious regional inequalities, which can be precisely identified from data on unequal distribution of regional income. Brazil's national wealth and economic centres are concentrated in the southeast region surrounding São Paulo municipality. The research by Azzoni (2001) suggests that São Paulo, which occupies less than 3 percent of total national land area, has accounted for about 36 percent of Brazil's total national domestic product since 1949. For instance, in 1996, income per head in the richest region, São Paulo was more than 6 times higher than in the poorest region, Piauí, at US\$ 1,063 and 6,547 respectively (2001, p. 135). Azzoni (2001) is thus able to illustrate that Brazil has suffered from sharp spatial disparity among its regions by empirically analyzing regional income distribution using the GDP measurement approach.

However, it is undoubtedly the case that the GDP approach has been the butt of much criticism (e.g. Fields, 1980; Sen 1992, 1999; Haq, 1999). One major criticism of the GDP lies in its inability to show regional income distribution. Rapid economic and

industrial development of a country (as demonstrated by income growth) does not necessarily reflect income distribution among individual regions. Sen (1999) argues that the issue of income distribution among individual people and regions is as important as GDP growth, whilst Fields (1980) claims that income distribution between the poor and rich cannot be presented solely in the growth rate analysis of GDP or industrial development. The poor might be not able to enjoy the benefits of economic development and GDP growth.

Economic development does not necessarily lead to reduction of regional income inequality: in fact, the opposite may be true. Certain models of economic growth might worsen regional wealth distribution and further widen regional inequality. In a later study, Fields (2001) was unable to find an absolute link between economic growth and decreasing income inequality. Field (2001, p. 65) states, "Nor is any empirical tendency found for the rise or fall in inequality to be linked to the rate of economic growth in a country or to the initial level of inequality." Numerous cases suggest that the correlation between economic development and regional inequality reduction is quite weak or even completely negative, but this in fact depends on the model of economic development. For example, governmental policy can directly affect wealth redistribution and balanced regional development. Regional disparity within some developed regions and nations, such as Taiwan (China) and South Korea, seems to be low and still falling. In contrast, regional income inequality tends to be quite high and continually increasing in some middle or low income countries, such as Brazil and Colombia (Field, 2001). Ahluwalia (1976) rightly argues that GDP per capita lacks the power to explain the functioning mechanisms by which economic growth influences inequality.

Secondly, many critics point out that the GDP approach can neither capture every aspect of human development nor measure disparities in areas such as public health and education. Citing the cases of Sri Lanka and Guinea, an overall weak correlation between economic growth and human development has been claimed by Haq (1999). He argues that rapid economic development does not necessarily lead to the improvement of human well-being; in addition, a high rate of economic growth does not directly link to development of human welfare. Although both countries have the same level of income per capita (US\$ 500), Sri Lanka has done much better than Guinea in terms of human development. Huge development inequalities between the two countries were found, in areas such as life expectancy, adult literacy and infant mortality rate (Haq, 1999). Haq firmly believes that,

It is not just the level of income that matters. It is how society spends that income. Also important are the many choices that human beings make - particularly in social, cultural and political areas - that may be largely independent of their income. The quality of growth is more important than quantity. (1999, p. 40)

By measuring the growth of income per head alone, it is not easy to analyze inequalities of education and public health among individual regions and people as income growth of a country does not necessarily transfer to developments in public services. This is especially true of the equal regional development of basic education and health, rather than income growth plays a crucial role in evaluation of overall

development growth. They cite the case of the sharp disparity in provision of basic primary education between India and South Korea. For example, in 1990, many gifted students from both countries gained PhD degrees from overseas institutions. However, India's provision of basic primary education was far more patchy and backward than South Korea's. In India, more than 50 percent of the population had never enrolled in the school or been educated: while the equivalent figure was less than 10 percent in South Korea.

The third prominent issue concerning many scholars is the fact that GDP fails to address the cost of environmental damage in calculating growth. Ray's study (1998) clearly points out the shortcomings of income approach in measuring social and the other external costs, such as pollution, human suffering caused by dams, railways and other major construction projects, and unsustainable natural resource consumption.

Environmental damage not only has a negative influence on economic growth, but also seriously affects public health and long-term human development. China is probably one of the best examples of this issue. Since 1978, China has made remarkable achievement in terms of fast economic and industrial development. Income per head has been increasing substantially over the past two decades, with more and more industrial enterprises being established. In addition, the various commodities, which range from steel, cement and aluminium to home-entertainment systems, have been increasingly exported to international markets during the 'reform and open-door' period. As far as the growth of total material wealth is concerned, China is undoubtedly a big gainer. However, Barister's study (1998) discusses the challenge of environmental pollution, including air, water and soil pollution, to public health and human development in this nation, whilst Rob Gifford (2007) suggests severe environmental problems, such as water shortage, have impacted negatively on human health.

However, unfortunately, the GDP approach cannot tell us anything about the environmental damage to China caused by its dramatic economic development. The environmental issue must be addressed in order to achieve sustainable economic growth and long-term enhancement of human welfare, and indeed, the new concept of 'Green GDP' has attracted worldwide attention in recent years. By adopting a green GDP calculation, environmental damage and costs caused by GDP growth can be empirically measured; the real quality of economic and industrial growth can be properly estimated. Green accounting would largely exempt natural costs and environmental damage in calculating traditional domestic products and income even if they contributed to the total amount of GDP. Boyd highlights the importance of green GDP in measuring overall human well-being and social development. As Boyd (2006, p. 9) points out,

... green GDP also allows period-to-period comparison of the quantity of ecosystem services over time ... Degradation or enhancement of services can be directly measured and reflected in the year's green GDP numbers.

Meanwhile significantly, the intention of using green GDP calculation in measuring environmental costs is not easy to achieve. Although China is the first country to attempt green national accounting, it has given up this new approach to calculating GDP due to various practical problems, including calculation method, technology,

lack of credible data and local authority resistance³ (Boyd, 2006; Chinese Government's Official Web Portal, accessed in 24th Feb 2008).

In addition, to environmental pollution, Hamilton (1998) argues that another important factor has been ignored by the GDP measurement approach – families' informal economic contribution, which is not included in market transactions. Hamilton (1998) cites the specific case of cultivating cash crop production by farmers. Hamilton's argument is shared by Scitovsky (1997), whose research suggests that GDP inequality estimation, which only measures economic activities under market conditions, is insufficient to evaluate overall national welfare because not all needs, expectations and desires of human beings are reflected in market transactions.

In conclusion, income approach is a fair and convenient way of measuring spatial economic disparity and the simplicity of GDP calculation makes it a popular measurement approach. However, it is reasonable to claim that this approach has some serious drawbacks, including its failure to deal with the distributional issue among individuals or to address the impact of environmental damage on a nation's health.

2.2.2 HDI Measurement Approach⁴: Advantages and Criticisms

According to the studies conducted by Sen (1992; 1999), Haq (1999) and the UNDP (1990), the Human Development Index includes three measurement indices with equal weighting – GDP per capita, life expectancy and education (as measured by both adult literacy and school enrolment rates). These studies suggest, that by measuring these three indices, disparity of human development opportunities can be properly estimated: as they capture inequalities in both economic and social welfare by including income, health and education. This reinforces the evidence of numerous previous studies that HDI is superior to GDP in measuring regional inequality and poverty due to its multidimensional approach. On the one hand, the HDI can highlight many important social aspects of spatial inequality. On the other hand, it can help researchers better evaluate regional disparity in terms of human development⁵. In addition, it augments our understanding of the real meaning of income growth and economic development.

³ In China, the new system of green national accounting was officially implemented in 2004 by a joint technical team from two Chinese governmental institutions – State Environmental Protection Administration and National Bureau of Statistics. The green accounting project covered 42 diverse industries and all of the eastern, central and western regions. However, in May 2006, China abandoned this green national accounting. (e.g. Chinese Government's Official Web Portal, accessed on 24th Feb 2008; Boyd, 2006)

⁴ "The HDI has three key components: longevity, knowledge and income. Longevity is measured by life expectancy at birth as the sole unadjusted indicator. Knowledge is measured by two education variables: adult literacy and mean years of schooling, with a weight of two-thirds to literacy and one-third to mean years of schooling." (Haq, 1999, p. 49)

⁵ According to the definition from the UNDP (2000, p. 2), "Human development, in turn, is a process of enhancing human capabilities – to expand choices and opportunities so that each person can lead a life of respect and value. When human development and human rights advance together, they reinforce one another – expanding people's capabilities and protecting their rights and fundamental freedoms."

Many academics who support the Human Development Index have cited worldwide cases to illustrate the limitations of GDP approach in measuring inequality and poverty. For example, research conducted by Sen (1992) on the issue of regional inequality in India, cites the case of an Indian state – Kerala to argue the discrepancy between measurement indicators of human and economic development. Among all the states within India, Kerala's development performance in social and human aspects, including life expectancy at birth, literacy and infant mortality rates, is outstanding. For example, relative to India's average figure of 57 years in India, the average life expectancy at birth in Kerala was above 70 years. In addition, the average literacy rate of Kerala was 91 percent: in contrast to just 52 percent in India as a whole. Indeed, in the "crucially important functionings", as Sen (1992, p. 126-27) called them, Kerala has done much better than many other Indian states in social welfare development, even though the income per head in Kerala is probably among the lowest in India (Sen, 1992).

Sen argues that successful development should be judged on development of freedom and human capabilities rather than GDP economic growth. Fast-developing countries with high income-levels might still face the challenge of disparities and poverty within their various regions, accompanied by social problems such as illiteracy, AIDS and crime. Although economic development in India is rapid; it has been facing the serious challenge of rising regional inequality. The study by Deaton and Drèze (2002) suggests a huge gap in economic development between the developed western and southern regions (Andhra Pradesh excluded) and the under-developed northern and eastern regions. More seriously, due to lack of government investment in education and public health care, the achievement of rapid economic growth has brought limited success for India in terms of human development. In general, Drèze and Sen (1995) conclude that the Indian government's efforts to enhance human development and reduce regional disparities are entirely unsatisfactory and the later research by Sen (1999) provides striking argument to reinforce these claims. Sen (1999, p. 100) states,

... but there are large areas within India where life expectancy and other basic living conditions are not very different from those prevailing in these most-deprived countries.

India's performance in female literacy education was even more disappointing. Drèze and Sen's study (1995) reveals that the rate of adult female literacy in China is far higher than in India. Moreover, even the overall performance of adult female literacy in the sub-Saharan African countries is better than India's. More than 50 percent of females in India, aged between 15 and 19 years, are illiterate. These striking figures cast doubt on whether the rapid economic growth in India during the past decade can be regarded as an example of development miracle. The writer is not convinced by this argument. India has been facing severe development disparities and poverty among the various regions and individuals. Undoubtedly, India has been facing severe development disparities and poverty at regional and individual levels and has a long way to go in order to overcome these problems.

According to the World Development Report (1995), in 1993, Pakistan's GDP per capita was almost the same as the equivalent figure in Sri Lanka. However, the infant mortality rate in Pakistan was much higher than that of Sri Lanka. In addition, the literacy rate of Pakistan was only about one-third of the figure in Sri Lanka. Moreover,

per capita income in South Africa and Brazil, of US\$ 2,470 and 2,540 respectively, is much higher than in China (US\$ 350) and Sri Lanka (US\$ 430). It could be assumed that human welfare development in Brazil and South Africa would be much better than in China and Sri Lanka. However, this is apparently not the case as, for example, the figures for life expectancy at birth in these two low-income countries were significantly higher than in South Africa and Brazil (Sen, 1992). Sen makes a similar comparison between Costa Rica and the USA. As shown in the human well-being indicators, Costa Rica is not much poorer than many developed European countries or the USA. Life expectancy in Costa Rica (75 years) was only one year less than in the USA (76 years), even though the latter's GNP per capita (US\$ 20,190) was more than 11 times higher than Costa Rica's (US\$ 1,780) (Sen, 1992).

During the period of fast economic development and high income levels, Brazil had faced problems of unequal human development and related social issues which are becoming characteristic of certain fast-growing economies (Drèze and Sen, 1995). Drèze and Sen (1995, p. 34) maintain that,

... the development experiences of some fast-growing countries during the last few decades has resembled one of 'unaimed opulence', combining high rates of economic growth with the persistence of widespread poverty, illiteracy, ill health, child labour, criminal violence, and related social failures.

2.2.2.1 Sen's Theory: Human Development Capabilities

Nobel Prize laureate – Amartya Sen, is probably one of the most respected researchers into human development capabilities, social welfare and human freedom. Sen is a productive writer and great thinker who has done important work on the issues of inequality, poverty and human development. Sen firmly argues that the GDP measurement approach cannot present all aspects of regional inequality and poverty within a country: especially in the vital area of human development. Sen (1992, p. 28) stresses that,

Such diversities [**in the human aspect**], however, can be hard to accommodate adequately in the usual evaluative framework of inequality assessment. As a consequence, this basic issue is often left substantially unaddressed in the evaluative literature. (Bold content added)

Sen's study proposes two crucial concepts for analyzing human development – capability and functionings⁶. Sen (1992, p. 83) offers a definition of capability, "Capability reflects a person's freedom to choose between alternative lives (functioning combinations), and its valuation need not presuppose unanimity regarding one specific set of objectives (or, as Rawls calls it, 'a particular comprehensive doctrine')." In his study, Sen (1992) discusses the sharp difference between human development and traditional income approaches in analyzing

⁶ Sen points out the relationship between capability and functionings, "Capability is, thus, a set of vectors of functionings, reflecting the person's freedom to lead one type of life or another" (1992, p. 40).

inequality and social development. He argues that the real difference between his human approach and the traditional GDP approach lies in the development of human capabilities: Sen refers to this as “capability approach”. Sen (1992) firmly believes that capability development can channel traditional material wealth into human functionings which make up the key factors for human development. The capability development of a person, which is referred to by Sen as the “capability set”, indicates the options for enlarging his/her human freedom and development potential. He seems to claim that human capability leads to the expansion of human development potential and freedom. Without capability development as measured by human functionings, human development is hard to achieve.

As previously stated, Sen (1992) believes that overall human inequality cannot be entirely evaluated by the GDP approach. Relative to a healthy person, he cites the case that a disabled person might face more personal difficulties in enlarging development potential, even if they have similar incomes. Sen argues that disability restricts freedom to enjoy human life and participate in various social and economic activities, and suggests that overall equality and social justice must be demonstrated through various aspects (e.g. education, health, political liberty) rather than income alone, and assessment of overall inequality should take into account these different dimensions of disparity.

Drèze and Sen (1995) claim that the fundamental goal for economic growth is to achieve human and social development, which is not entirely dependent on economic development, as measured by GDP growth, but also on many other important social and political factors. However, in a later study, Sen also expresses his understanding of the significance of wealth in affecting living standards and human development,

It is as important to recognize the crucial role of wealth in determining living conditions and the quality of life as it is to understand the qualified and contingent nature of this relationship. (1999, p. 14)

To sum up, Sen (1999) identifies the limitations of the traditional GDP approach in measuring human freedom and development capabilities, which are due to its inability to assess human development opportunities. Income growth and economic prosperity cannot alone guarantee development of human well-being.

Amartya Sen’s human development research has been given credit by many scholars (e.g. Kiron, 1997; Hicks, 2002; Haq, 1999) who believe that it is preferable to the insufficient and narrow measurement of the traditional GDP approach. Sen’s theory has not just been useful for estimation of the ultimate aim of economic growth, but more significantly, his research has offered an important framework for analyzing social and human development. Kiron (1997) claims that Sen has built a fundamental theoretical framework for deep analysis of multidimensional social and human welfare. Hicks (2002) also supports Sen’s achievement in broadening the discussion on human and social development and believes that Sen’s concepts of human capability and functionings allow researchers to capture the full spectrum of human development. As Hicks (2002, p. 139) precisely states, “One of the principal innovations of Sen’s capability approach is to widen the scope of what counts in the determination of societal development and individuals’ well-being.”

Which is the ideal approach to estimation of human development and overall national welfare? Some scholars argue that the HDI is the only multidimensional measurement approach which can evaluate both economic growth and human development, altering the focus of regional income inequality and putting human development back at centre stage⁷. They highlight that human development and increased freedom are the ends, not the means, of economic development and increasing income is not the fundamental goal of economic growth, but a means to achieve human freedom and strengthen human development capabilities. Therefore, as maintained by Streeten (1994), the income-oriented GDP approach is hardly the best choice to measure inequality and poverty.

2.2.2.2 Haq's Theory: Human Development

Haq is another prominent researcher into the issue of human development and inequality measurement. He argues that both the policymakers and academics seem to have forgotten the fundamental focus of economic development: the people. How should we judge the success of economic development? Haq (1999) points out that the enlargement of human development choices is the basic objective of economic and income growth. The judgement of successful economic development lies with the people: whether or not their living standards, health and environment, can be improved. He believes that personal income is crucial to development of human potential and freedom: but wealth is not the only necessary factor. Haq suggests that a democratic country does not need to be a rich one; a country with free elections, rule of law and a democratic constitution could be quite poor. In addition, the gender equality within a nation is unrelated to distribution of GDP or income growth. Haq (1999, p. 15) further maintains that,

Many human choices extend far beyond economic well-being. ... The use that people make of their wealth, not the wealth itself, is decisive.

Haq (1999) strongly criticizes the GDP approach to measuring inequality due to its inherently ill-conceived definition and omission of data from some informal economic activities outside the market economy. Moreover, Haq argues that the GDP approach is unable to measure income distribution among individual households and regions. Haq invented and actively promoted the Human Development Index for measuring regional inequality, poverty and human development and believes that it can effectively measure the many strands of social welfare. By adopting this measurement approach, policymakers can better understand the fundamental objective of economic growth rather than its means alone. Moreover, Haq (1999) points out the advantages of HDI over the income approach in inequality measurement; for example, more meaningful regional comparison by avoiding extreme situations, and demonstration of the general situation across the whole society.

⁷ The initial promotion of human development and freedom by scholars can historically be traced back a long way. Haq (1999, p. 13) states that "The rediscovery of human development is not a new invention. It is a tribute to the early leaders of political and economic thought. The idea that social arrangements must be judged by the extent to which they promote 'human good' dates back at least to Aristotle (384 – 322 B. C.)."

2.2.3 Comparison and Evaluation: GDP and HDI

It is justifiable to argue that the income approach has certain limitations in analyzing regional development and measuring human inequalities. One of the big problems Todaro and Smith's study (2006) suggests, regarding GDP growth, is its failure to assess income distribution among individual households and regions. For example, Oman's income per head (US\$ 6,700) was more than 13 times higher than the equivalent figure in Sri Lanka (US\$ 500). However, in terms of human development, Oman's position is inferior to Sri Lanka's: as is reflected in the fact that Oman's life expectancy at birth (66 years) was significantly lower than Sri Lanka's (71 years) (Streeten, 1994).

From my point of view, the Human Development Index developed by Haq and Sen and their theories of human development capabilities have broadened our knowledge of human development. The income approach now appears to be inadequate for analyzing regional inequality and human poverty as it is important to estimate the impact of human well-being development on overall development process. The HDI can help us to properly understand the key meaning and fundamental goal of economic development. The UNDP (2000) suggests that the fundamental aim of economic development is to achieve human development; and from this aspects, Sen's thought on human development capabilities and his associated HDI approach deserve to be taken seriously. Even though the ecological issues have not been addressed by this index, Sagar and Najam (1998) point out that the HDI offers a better discussion of the real meaning of human development than the income approach. More significantly, as Haq's research (1999) details, the income approach has serious shortcomings in terms of social and human welfare issues.

One of the key arguments made by many pro-HDI scholars is that the GDP approach leads to inappropriate judgements on the objectives of economic development as it over-emphasizes economic development as measured by income growth, and deliberately ignores the other important aspects of social disparity. Sen (1999, p. 14) stresses that,

The ends and means of development require examination and scrutiny for a fuller understanding of the development process; it is simply not adequate to take as our basic objective just the maximization of income or wealth ...

As far as human development is concerned, economic development and income growth are not everything. The GDP growth cannot ensure development of human well-being and other social aspects. This is precisely why Ray (1998, p. 8) suggests, "However, a high and equally accessible level of material well-being is probably a prerequisite for most other kinds of advancement, quite apart from being a worthy goal in itself." In particular, government policy has the power to extend the benefits of economic development among the various regions and individual people. Policies of economic development and public investment, implemented by the government, would have a positive effect on human development. Haq (1999); and Drèze and Sen (1995) also stress the important role played by the government in dealing with inequality and poverty.

However, some important points need to be stressed here. It is unfair to argue that the correlation between economic growth and human development is weak or even negative, due to the impact on human development of other non-income factors. It is also inappropriate to claim that income disparity and overall inequality only maintain a causal, but not necessary, relationship. For example, Sen (1999, p. 108) points out that the link between income disparity and social development inequalities is quite weak and states,

Empirically, the relationship between income inequality and inequality in other relevant spaces can be rather distant and contingent ...

I strongly question this view. Some parts of the argument made by Sen are problematic and his case study is unconvincing. My view is this: the relationships between various inequalities across economic, social, political and human spheres are strongly correlated. Income inequality is the fundamental cause of disparities in other social aspects. Although income equality would not necessarily transfer to improvement of human well-being or reduce disparities, it makes up the most important development component of overall economic, social and human equality. Income equality lays the key foundation for correcting the worsening trend of unequal human development and unbalanced regional growth. In fact, in an early study by Arthur Lewis (1955), he was already suggesting the crucial role of economic growth in contributing to human development and increasing human potential. Moreover, the research conducted by Sagar and Najam (1998) suggests that high income has played a key role in expanding human choices and improving standards of living. In addition, Fields (1980) highlights that income is the best measurement indicator for economic growth and general development. Similarly, Debraj Ray (1998, p. 170) points out that,

At the level of philosophy, the notion of inequality can dissolve into an endless sequence of semantic issues. Ultimately, economic inequality is the fundamental disparity that permits one individual certain material choices, while denying another individual those very same choices.

Interestingly, one of the key conclusions reached by Sen is that GDP growth by itself is not sufficient to stimulate human well-being improvement and enlarge human development freedom; however, even Sen (1992) himself admits the importance of low income in causing poverty and human development disparity, as he says,

It can be argued that poverty is not a matter of low well-being, but of the inability to pursue well-being precisely because of the lack of economic means. (1992, p. 110)

Economic development as mainly reflected in income growth, is absolutely essential for improving the welfare of the poor. Even Haq (1999) has admitted that economic growth should not be rejected as sustainable human welfare development inherently requires growth. In addition, Human Development Report 2000 published by the UNDP (2000, p. 9) suggests, "Poor countries need faster growth to generate the resources to finance the eradication of poverty and the realization of human rights."

Only when achieving rapid economic development and associated high-income levels as shown in GDP per head, can the government spend more money on improving public health and education. Moreover, only then does a country have sufficient financial capability to improve human welfare and expand freedom of human development. It is important to note: human development is desirable while economic growth is also essential and both human and economic development need to be taken seriously. Economic growth is the basic requirement for human development and the two should not be separated.

In my opinion, some part of Haq's argument are unjustifiable and highly problematic. He seems to underestimate the role played by income growth and economic development in ensuring human development and social welfare. He seems not to fully recognize that economic growth still holds the best cards for human welfare development. Without economic growth, the enlargement of human development choices, desired by Haq, in the form of cultural development, entertainment, good living environment and political participation, could not be pursued. The realization of free choice in human development requires material and financial resources. In this respect, Haq's argument is not realistic; it is too strong and subjective. Haq (1999, p. 222) inappropriately claims, "Gone are the days when the prosperity of a society was identified only with the measure and growth of its GNP."

The evidence of Sri Lanka, Cuba, and China during the pre-reform period illustrates that human welfare improvement, as reflected in the rise of life expectancy at birth, can be achieved, even without dramatic increase in income per capita. It might be assumed that human development does not necessarily relate to economic development. However, this view is not appropriate. All of these countries made economic growth during the period of human development even though the growth of GDP per capita was fairly small. More significantly, the income distribution among individual people in these nations was relatively equal. In China, without question, personal income distribution was far more equal during the pre-reform period than the reform decades since 1978. Therefore, these cases do not support the view that income growth and economic development have nothing to do with human development.

Is the GDP really out of date as a measure of spatial inequality? Is traditional income growth theory incapable of reflecting human development in terms of extending human capability and freedom? Is the HDI measurement completely different from the income approach? Is the HDI a valid alternative approach to inequality and poverty measurement? The answers to these questions are probably negative. Sugden (1993) raises questions of his own on Sen's human development framework.

Undoubtedly, with rapid development of economy and technology, people are getting more information about human and social development; and their understanding on the importance of full-range human development has been significantly enhanced. It is reasonable to argue that people expect economic development to be more responsive to their personal needs and to focus on human and social welfare. It is undeniable that human development should be highlighted as the central purpose of development. However, this does not necessarily mean discrediting the important role played by the GDP and income growth on the long-term development process. Due to their various shortcomings and inherent problems of biased equal component

weighting, human disparity indicators like the HDI are hardly desirable approaches to measurement of human development disparity and regional inequality, whereas the income approach is still a fair method and has a crucial role to play. The GDP is perhaps one of the best and most popular approaches in assessing human development disparity and overall social welfare as economic development and income growth are the key to improvement of human well-being.

Even though there are some discrepancies between income and human well-being indicators, Streeten's study (1994) suggests, in a worldwide comparison of countries, that the relationship between the various indicators of economic and human development is highly positive. Moreover, empirical study of certain developing nations made by Ray (1998) confirms the highly positive correlation between income and other aspects of human development. In fact, Sen (1999, p. 72) himself admits that,

The view that poverty is simply shortage of income is fairly well established in the literature on the subject. It is not a silly view, since income - properly defined - has an enormous influence on what we can or cannot do.

During the post-World War II period, as suggested by Ray (1998), many East Asian countries have made significant progress in social development as demonstrated in the high enrolment rate of primary and secondary education in countries such as Singapore and South Korea. Meanwhile significantly, in my estimation, it should not be forgotten that during the same period, many East Asian countries have also achieved rapid economic growth and income per head has been widely discussed by academic scholars, politicians and media. As stressed by Harris (2000), the fact that high economic growth leads to human development should not be underestimated or, worse still, ignored.

In terms of the sub-Saharan African continent, the causality between slow economic growth, worsening human inequality and restricted human development potential, is evident. In Africa, the underdevelopment of its economies and decreasing income per capita have accompanied slow human development and other serious social challenges. According to the figures presented in the Human Development Report 2007/2008 (2008), in terms of countries with low human development, the African countries accounted for the majority. During the period between 1990 and 2005, the annual growth rate of GDP per head in African countries with human underdevelopment was slow. Economic growth rates in the majority of these African countries were even below zero; this is true of Congo and Burundi, whose annual growth rate of GNP per capita were -5.2 and -2.8 percent respectively during the period between 1990 and 2005. It suggests that the economically backward position of Africa in the world has not changed: in fact, the African national economies are falling further behind. More significantly, during the same period, the situation of human underdevelopment in these African nations was severe, as demonstrated in social development indicators. In terms of public health, for example, in 2005, the prevalence of tuberculosis cases per 100,000 population was quite high compared to the rest of the world. The equivalent figures in Sierra Leone and Togo were as high as 905 and 753 respectively. In many African countries with low human development, the number of adults living with HIV/AIDS was striking. In Zambia and Malawi, for

instance, for people aged between 15 and 49, the HIV infection rates were 17 and 14.1 respectively. In addition, as far as education is concerned, there were still significant numbers of adult who were completely illiterate and the adult literacy rates in Mali and Chad were merely 24 and 25.7 (See Table 2.1 below). The striking figures suggest that economic growth and human well-being development are highly correlated. Economic development and social welfare improvement cannot be deliberately separated for the purpose of academic analysis as economic growth accounts for human development. The cases from Africa and other underdeveloped countries indisputably illustrate that slow economic and income growth is the key factor responsible for human welfare underdevelopment and disparity.

Table 2.1: Human and Economic Development in Selected African Countries with Low Human Development

Country	Human Development Indicators			Economic Growth Indicator
	Tuberculosis Cases ^c (per 100,000 people) 2005	HIV prevalence ^a (% aged 15-49) 2005	Adult literacy rate (% aged 15 and above) 1995-2005 ^d	
Sudan	400	1.6 [0.8-2.7]	60.9 ^m	3.5
Togo	753	3.2 [1.9-4.7]	53.2	n.a.
Nigeria	536	3.9 [2.3-5.6]	69.1 ^h	0.8
Congo, Dem. Rep. of the	541	3.2 [1.8-4.9]	67.2	-5.2
Zambia	618	17 [15.9-18.1]	68	-0.3
Senegal	466	0.9 [0.4-1.5]	39.3	1.2
Gambia	352	2.4 [1.2-4.1]	n.a.	0.1
Malawi	518	14.1 [6.9-21.4]	64.1	1
Rwanda	673	3.1 [2.9-3.2]	64.9	0.1
Mali	578	1.7 [1.3-2.1]	24	2.2
Central African Republic	483	10.7 [4.5-17.2]	48.6	-0.6
Chad	495	3.5 [1.7-6.0]	25.7	1.7
Burundi	602	3.3 [2.7-3.8]	59.3	-2.8
Niger	294	1.1 [0.5-1.9]	28.7	-0.5
Sierra Leone	905	1.6 [0.9-2.4]	34.8	-1.4
Average	547.6	4.75	50.6	-0.01

Note:

^a Data are point and range estimates based on new estimation models developed by UNAIDS. Range estimates are presented in square brackets.

^c Data refer to all forms of tuberculosis.

^d Data refer to national literacy estimates from censuses or surveys conducted between 1995 and 2005, unless otherwise specified. Due to differences in methodology and timeliness of underlying data, comparisons across countries and over time should be made with caution.

^h UNESCO Institute for Statistics estimates based on its Global Age-specific Literacy Projections model, April 2007.

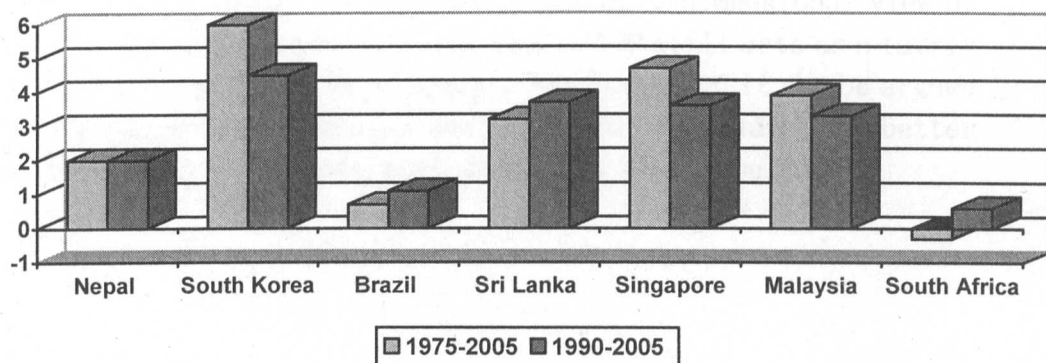
^m Data refer to North Sudan only.

Source: UNDP (2008, pp. 258-60; pp. 270-72; pp. 278-80)

Citing cases such as Brazil, Costa Rica and Sri Lanka, many pro-HDI scholars have strongly criticized the limitations of GDP approach in measuring full-range development; as was discussed earlier in this chapter, Drèze and Sen (1995) argue that development is a complex and comprehensive framework. The growth of GDP does not indicate overall improvement of national welfare; moreover, it does not guarantee increases in human freedom or human capabilities. For example, Harris's study (2000) suggests a dramatic difference between the HDI and GDP measurements in Costa Rica, Sri Lanka, Brazil and Pakistan. However, the arguments made by many pro-HDI scholars are only partially correct. It is certainly the case that Sri Lanka has done better in terms of human development than many other fast-developing countries, such as Brazil, which have achieved average high-income levels in worldwide comparisons. Nevertheless, another side of the story should not be ignored. The economic growth rate in Sri Lanka was much higher than that of Brazil during the period between 1975 and 2005 even though the absolute income figure in Sri Lanka was still quite low. In the period of 1975-2005, Sri Lanka's annual growth rate of GDP was 3.2 percent, while the equivalent figure was merely about 0.7 percent in Brazil. In the period between 1990 and 2005, Brazil's average growth rate of GDP was merely 1.1 percent: in contrast, the equivalent figure was 3.7 percent in Sri Lanka (see Table 2.2 below). Therefore, its relatively fast economic growth might enable Sri Lanka to perform better than Brazil in terms of human development by putting more public investment into health, hospital construction, basic education and other social welfare. This possibility simply can not be eliminated. Without the foundation of rapid economic growth, Sri Lanka's superior performance in human development would have been difficult to achieve.

This is also valid when we analyze other countries' in terms of human development; for example, South Korea – South Africa and Nepal – Singapore. According to the figures presented in Table 2.2, the average economic growth in both South Korea and Singapore were much higher than the equivalent figures in Nepal and South Africa during the period from 1975 to 2005. For example, the annual growth rates of income per capita in South Korea and South Africa were 6.0 and -0.3 percent respectively. In addition, relative to the equivalent figure 4.7 in Singapore between 1975 and 2005, the annual growth rate of GDP per capita was only 2 percent in Nepal. These figures appear to reflect the dramatic human development in South Korea and Singapore in contrast to the poor performance of social welfare in South Africa and Nepal. In terms of the East European countries, Haq (1999) questions the likelihood of increasing trends of the HDI in these nations due to their slow economic development. Development strategy related to high-growth is of course not adequate, but this does not justify complete rejection of economic development and income growth. Moreover, we should not be lead to another extreme – a discussion that ignores economic growth. Without economic growth, human development is simply unachievable.

Table 2.2: Annual Growth Rate of GDP Per Capita in Selected Countries 1975-2005 (%)



UNDP (2008, pp. 277-280)

Data published by the Human Development Report 2007/2008 (2008) demonstrates the close relation between human development, life expectancy and economic growth. Income growth has played an important role in contributing to overall human development. In terms of the regions with high and medium human development (0.897 and 0.698), the average GDP per capita were US\$ 23,986 and 4,876 in 2005. By contrast, the income per capita was merely US\$ 1,112 for the region with low human development (0.436). For the regions with high income (US\$ 33,082), the average HDI value was 0.936 in 2005, which was more than 64 percent higher than the equivalent figure in the low income region (0.57). (2008, p. 232)

The figures presented in Table 2.3 below illustrate the changing trends both for HDI and per capita in selected countries across the world. These countries represent high, medium and low levels of human development. Some important findings have been generated in the comparison analysis. On the one hand, the economic growth rates in the five countries (Norway, Iceland, Australia, Canada and United States), which have the highest average levels of human development, have increased significantly, as demonstrated in the rising trends of GDP per head. The steadily increasing HDI in these top five countries is accompanied by fast growth of per capita GDP. Singapore and South Korea demonstrate similarly close correlation between HDI and GDP measurement. In terms of the countries with medium human development, such as Thailand, Sri Lanka and China, the rising figures of GDP also illustrated the dramatic upward trend of HDI. For example, Sri Lanka's Human Development Index rating had increased more than 21 percent to 0.743 between 1975 and 2005. However, during the same period, GDP per head in Sri Lanka reached PPP US\$ 4595 in 2005 from only US\$ 382 in 1975. Countries with low human development presents a completely different overall picture. Selected African countries such as Congo, Zambia, Burundi and Niger had shown a rapidly decreasing trend in income per head. Importantly, the HDI performance in these five African countries was also entirely unsatisfactory, either falling (e.g. Zambia and Congo) or increasing slightly (e.g. Niger).

In general, the comparison analysis between the countries and regions with the various development levels of human well-being indicates the close relationship between economic growth and human development. The influential role played by economic growth in contributing to human development and freedom is notable. As Ray (1998, p. 29-30) states,

It is arguable that although taking a wider and multidimensional view of development is conceptually correct, per capita GDP still acts as a fairly good proxy for most aspects of development. For instance, it can be argued that rising income levels ultimately and inevitably translate into better health, nutritional, and educational standards in a population.

Table 2.3: Changing Trends for Human Development Index and Per Capita Income in Selected Countries 1975-2005

Country	Human Development Index (HDI)					GDP per capita (1995 US\$)						
	1975	1980	1985	1990	1998	2005	1975	1980	1985	1990	1998	2005 (PPP US\$)
High Human Development												
Norway	0.853	0.872	0.883	0.895	0.934	0.968	19,022	23,595	27,113	28,840	36,806	41,420 ^f
Iceland	0.857	0.879	0.888	0.906	0.927	0.968	17,445	22,609	23,977	26,510	29,488	36,510
Australia	0.841	0.858	0.87	0.884	0.929	0.962	14,317	15,721	17,078	18,023	21,881	31,794
Canada	0.865	0.88	0.902	0.925	0.935	0.961	14,535	16,423	17,850	19,160	20,458	33,375
United States	0.862	0.882	0.894	0.909	0.929	0.951	19,364	21,529	23,200	25,363	29,683	41,890 ^f
Singapore	0.725	0.756	0.785	0.823	0.881	0.922	8,722	11,709	14,532	19,967	31,139	29,663
South Korea	0.684	0.722	0.765	0.807	0.854	0.921	2,894	3,766	5,190	7,967	11,123	22,029
Costa Rica	0.732	0.756	0.756	0.775	0.797	0.846	2,231	2,482	2,176	2,403	2,800	10,180 ⁿ
Malaysia	0.62	0.663	0.696	0.725	0.772	0.811	1,750	2,348	2,644	3,164	4,251	10,882
Average	0.782	0.808	0.827	0.85	0.884	0.923	11,142	13,354	14,862	16,822	20,848	28,638
Medium Human Development												
Thailand	0.6	0.643	0.673	0.708	0.745	0.781	863	1,121	1,335	2,006	2,593	8,677
Sri Lanka	0.612	0.648	0.676	0.699	0.733	0.743	382	452	536	590	802	4,595
China	0.518	0.548	0.584	0.619	0.706	0.777	138	168	261	349	727	6,757 ⁿ
Average	0.577	0.613	0.644	0.675	0.728	0.767	461	580	711	982	1374	6,676
Low Human Development												
Zambia	0.444	0.456	0.47	0.451	0.42	0.434	641	551	483	450	388	1,023
Burundi	0.281	0.306	0.334	0.339	0.321	0.413	162	176	198	206	147	699 ⁿ
Congo, Dem. Rep. of the	0.416	0.43	0.447	0.45	0.43	0.411	392	313	293	247	127	714 ⁿ

Central African Republic	0.332	0.35	0.371	0.372	0.371	0.384	454	417	410	363	341	1,224 ⁿ
Niger	0.236	0.259	0.257	0.273	0.293	0.374	298	328	242	235	215	781 ⁿ
Average	0.341	0.36	0.375	0.377	0.367	0.403	389	357	325	300	243	888

Note:

^f For purposes of calculating the HDI, a value of 40,000 (PPP US\$) was applied.

ⁿ World Bank estimate based on regression.

^u World Bank estimate based on a bilateral comparison between China and the United States (Ruoan and Kai 1995).

Source:

1. UNDP (2000, pp. 178-81)
2. UNDP (2008, pp. 229-32)

As far as China is concerned, in general, the correlation between the GDP and HDI approaches in measuring regional inequality is positive (see the specific figures presented in Table 2.4). According to the empirical estimation presented in the UNDP China Human Development Report 2005 (2005) except for Guangxi, Qinghai and Xizang which illustrate wide variation between HDI and GDP, the rankings of HDI and income per capita in the other Chinese provinces were generally similar in 2003. In about one third of the Chinese provinces, both HDI and GDP rankings are almost the same. In addition, in another 50 percent of the Chinese provinces, a small variation between the HDI and income per head has been demonstrated. Khan and Riskin (2001, p. 148) point out that,

Human development, as measured by the Human Development Index (HDI) estimates for China's provinces, closely follows provincial per capita income.

Table 2.4: Rank of GDP and Human Development Indices in China's Provinces 2003

Province	HDI Rank	GDP Rank	GDP Rank Minus HDI Rank	GDP Index	Human Development Index (HDI)
Shanghai	1	1	0	0.919	0.909
Beijing	2	2	0	0.856	0.882
Tianjin	3	3	0	0.824	0.855
Zhejiang	4	4	0	0.778	0.817
Liaoning	5	8	3	0.721	0.808
Guangdong	6	5	-1	0.752	0.807
Jiangsu	7	6	-1	0.748	0.805
Heilongjiang	8	10	2	0.686	0.786
Fujian	9	7	-2	0.729	0.784
Jilin	10	13	3	0.650	0.776
Shandong	11	9	-2	0.714	0.776
Hebei	12	11	-1	0.670	0.766
Hainan	13	16	3	0.631	0.761
Xinjiang	14	12	-2	0.656	0.757
Hubei	15	14	-1	0.644	0.755
Shanxi	16	19	3	0.612	0.753
Hunan	17	18	1	0.615	0.751
Chongqing	18	21	3	0.607	0.745
Henan	19	17	-2	0.615	0.741
Inner Mongolia	20	15	-5	0.643	0.738
Jiangxi	21	24	3	0.594	0.732
Guangxi	22	28	6	0.575	0.731
Shaanxi	23	25	2	0.589	0.729
Sichuan	24	27	3	0.587	0.728

Anhui	25	26	1	0.588	0.727
Ningxia	26	23	-3	0.594	0.712
Qinghai	27	20	-7	0.608	0.684
Gansu	28	30	2	0.547	0.675
Yunnan	29	29	0	0.567	0.657
Guizhou	30	31	1	0.491	0.639
Xizang (Tibet)	31	22	-9	0.599	0.586

UNDP (2005, pp.154; pp.160)

In addition, The Human Development Report 2005 (2005) highlights that the provinces with high levels of HDI are mainly located in the wealthy eastern region; by contrast, most of the provinces located in the interior western region have recorded notably low HDI levels.

Numerous worldwide cases discussed earlier suggest why the arguments made by Streeten (1994) and other pro-HDI scholars are not so convincing. Streeten (1994) inappropriately suggests that large inequalities would be impossible to show, or even be entirely ignored by the GDP measurement approach. As Streeten (1994, p. 235) states, "As has been seen, one of the great drawbacks of average income per head is that it is an average that can conceal great inequalities." My view is this: no matter how inadequate the GDP approach in measuring inequality and poverty, it is important to recognize that the GDP is one of the most reliable measurement approaches available. This view has been endorsed by many scholars, including Hamilton (1998). In addition, its inadequacy in dealing with the economic activities occurring outside market transactions, does not seriously challenge the overall accuracy of the GDP approach. With the development of modern market economy, economic contributions made by the family and other informal sectors, are continually shrinking: as pointed out by Fields (1989) and Mankiw (2006).

Due to the long-term practice of centralised economic planning, China still has economic activities and informal sectors which are not included in the market production system. However, the GDP approach does provide a fairly good indicator of regional economic disparity and general performance of economic development. Selden's study (1993) supported this view and pointed out that,

Nevertheless, even in a society such as China's ... and the roles of state and collective are powerful, income distribution provides one vital and measurable gauge of inequality and of the overall performance of the system. (1993, p. 138)

Although economic development might not automatically transfer to full-range human development and reduction of overall regional disparity, economic growth is absolutely essential for human and other aspects of social development. Economic growth is inherently good for human development, particularly for the masses of poor

and disadvantaged people: who without economic growth would have no chance to improve their lives. Quah (2002) provides, detailed discussion in support of this viewpoint.

Can the HDI fully measure regional inequality and human development? Is the HDI an ideal measurement indicator? Personally, I strongly dispute the claim made by many academics, that all disparities, including social, economic, political and human factors, can be entirely captured by the HDI. This index is not all-encompassing and cannot represent the numerous strands of social welfare and human development. Streeten (1994) stresses that human development is such a complex issue that it cannot be fully estimated by a single measurement index. Furthermore, Haq seems to make the inappropriate claim that the component indices adopted in the HDI approach are perfect for measuring overall human development. Haq (1999, p. 47) alleges that, "Several other variables were considered and discarded. They showed a significant correlation with the variables already chosen ..."

Those scholars including Sen, who favour the HDI approach, seem to be too confident with the representiveness and credibility of the three chosen equal inequality indices: income per head, life expectancy at birth and adult literacy rate. Their arguments are apparently over-subjective and irrational. How can these three indices stand for all human and social indicators of inequity? Where is the hard evidence to support this argument? Significantly, like the income approach, it should be realised that the HDI approach has failed to address the distributional issue and is unable to measure the distribution of human development benefits among individual people. In addition, the HDI approach has suffered from the limitations of its three disparity components and their equal weighting and a number of scholars have raised their concerns (e.g. Kelley, 1991; Srinivason, 1994; Sagar and Najam, 1998). The solid justification for equal indicator weighting in the HDI approach is as yet unproven and the arguments made by many pro-HDI scholars are ambiguous and problematic. Like Sen, many pro-HDI scholars have clearly failed to address the issues and the Human Development Report published by the UNDP also failed to systematically discuss the issue of equal component weighting.

Kelley (1991) criticizes the arbitrary decisions on equal indicator weighting in the HDI and lack of reasonable explanation by pro-HDI scholars on this issue; whilst Ray also expresses his concerns over the issue of weighting. Ray (1998, p. 43) states that,

We noted that just because an overall index is provided does not mean it should be necessarily taken seriously: the weights are, of course, quite arbitrary.

Life expectancy rose substantially in Sierra Leone during the period between 1960 and 1987. Although this is apparently an important indicator of development of human well-being in this country, it fails to show up in the national HDI due to the equal weighting of indicators. The ranking of HDI in Sierra Leone has been remained

unchanged during the same period (Kelley, 1991). Furthermore, Sagar and Najam (1998) strongly criticise the artificial adjustment of GDP in the HDI measurement approach. They cite the cases of Switzerland and Mexico. According to the estimation made by UNDP's report, both countries had similar 're-estimated' GDPs per head (a discrepancy of merely 3 percent); however, in reality, it is hard to imagine that these two countries have similar levels of economic development and human living standards. In fact, the opposite is true: economic development in Mexico is significantly lower than in Switzerland. In general, Sagar and Najam (1998) argue that the ranking of human development for many nations might be much lower than the HDI figures claimed by the UNDP and that the actual disparity of human development between the developed and under-developed countries could be much higher.

Income per capita and school enrolments are not of similar importance in measuring regional inequality; and equal weighting of these two indices in the HDI approach is an oversimplification. The importance of the income indicator should be highlighted in the HDI approach by giving more weighting for GDP per capita. However, it is important to clarify here that it is not my intention to claim that education is less important than economic growth. Education is of course crucial for development, but, I don't agree with Sen and Haq's claim that school enrolment rate is the best indicator for assessing education disparity.

The representativeness of the three selected measurement components of the HDI is also uncertain. For example, due to the variation in school quality and academic years among different nations, Srinivason (1994) argues that their statistical figures of school enrolment make meaningful comparisons difficult. Srinivason further stresses that national policymakers can hardly gain meaningful guidance from the measurement of HDI due to its inherent bias and weakness in theoretical framework and empirical analysis.

Moreover, in terms of the public health index, for people living in the deprived regions, a long life may not necessarily be as enjoyable as for people residing in the metropolis: as personal happiness does not solely depend on longevity. For example, people living in developed regions appear, in general, to have more cultural entertainment choices than those in isolated regions. Therefore, in my view, the measurement of human health should be also judged by other factors rather than longevity, which is neither the only nor probably the best indicator of health disparity measurement.

This writer does not think these three selected indices can fully measure the numerous dimensions of human development. For example, for any country, social harmony and political stability are crucial to long-term economic growth and human welfare development. History has demonstrated that, without social harmony and a peaceful environment, it is very difficult to pursue stable economic growth and human development. Social harmony plays an important role in developing human happiness

and Chan (1963) and Fung (1983) (cited in Chang, 2001) focus on the tight correlation between the two.

As an example, in Iraq, serious sectarian violence and social conflict have blocked the possibility of developing the national economy and social welfare. Among the civilian population, life itself can not be guaranteed, let alone the enlargement of human development potential. Similar cases can also be found in many poor sub-Saharan African countries, such as Rwanda, Congo Republic, Sierra Leone, and Estonia. These countries have been suffered from civil war or social violence for many decades. The latest research report conducted by the Asian Development Bank (2007) has indicated the close correlation between economic growth and social harmony, and the impact of rising inequality on social harmony and stability. As the report published by Asian Development Bank (2007, p. 9) suggests,

... high levels of inequality may lead to pressures to redistribute. ... Alternatively, the process of bargaining that accompanies the call for redistribution, ranging from peaceful but prolonged street demonstrations all the way to violent civil war, may be extremely costly.

Social harmony plays an important role in ensuring human development. However, none of these three chosen measurement indices adopted in the HDI approach can effectively reflect social harmony or political stability. Adopting an indicator to analyze social harmony, for example, measurement of annual cases of local social conflict⁸, referred to as the “social harmony index” by this study, is probably more desirable than HDI in measuring social harmony within a country. However, this chapter emphasizes discussion of inequality measures and is not the appropriate place to further address the issue of social harmony.

Furthermore, the increasingly serious issues of environmental and global climate change, democracy and political participation as measured by political freedom, are totally ignored by the HDI approach. The research conducted by Sagar and Najam (1998) suggests that the HDI demonstrates a problematic estimation of worldwide inequality. They stress that the issue of environment has been entirely unaddressed by HDI, especially the effect of environmental damage on human being⁹. In addition, the study by Kelley (1991) raises the drawbacks of HDI caused by under-measurement of human political rights, in the belief that these factors have a big impact on human welfare improvement, including human freedom and rights. In addition, Haq’s study

⁸Social harmony measurement can be reflected by incidence of conflict within a society. The research by Leung, Koch and Lu (2002, p. 202) suggests this, “Harmony is a concept with a long history in Confucian countries, and so are traditions of conflict avoidance. Many customs developed that reinforce the idea of a harmonious group as one where conflict is avoided.”

⁹ The damage of environment pollution has a serious impact on the living standards of human being and on human development potential. The research conducted by Harris (2000) highlights the significance of ecological and environmental factors in affecting human development.

(1999) highlights the failure of HDI to measure the important issues of environment and political participation. In fact, even Haq admits that the HDI has not yet covered all aspects of social and human development. Human development is a rich and comprehensive concept which is very difficult to reflect in its entirety in the HDI. As he states,

Several critics have pointed out that the variables included in the HDI are very limited. ... *This criticism is justified ...* (1999, p. 57-58)

It is best to recognize that the HDI will remain a partial reflection of reality. (1999, p. 58)

Due to varying developmental levels, inherited culture and values, and unique national characteristics, in practice, the subjectively-designed and inflexible HDI is very difficult to apply uniformly to different countries. Haq (1999) argues that countries at differing stages of development have their own unique characteristics and HDI might not be able to objectively illustrate such disparities. Moreover, due to the lack of widely-accepted standardised measurement of development performance among different countries, Srinivason (1994, p. 240) highlights that the HDI approach with its problematic measurement indicators is simply not realistic or functional. The research by Srinivason (1994, p. 240) suggests that,

The components of HDI, namely, life expectancy and educational attainment, are "functionings" in the Sen sense but their relative values need not be the same across individuals, countries, and socioeconomic groups. Besides, the "intrinsic" value of a single "functioning," namely, ability to live a healthy life, is not captured by its linear deprivation measure in HDI, since a unit decrease in the deprivation in life expectancy at an initial life expectancy of, say, 40 years is not commensurate with the same unit decrease at 60 years.

In my opinion, the single indicator is the most clear and straight-forward approach to inequality measurement, as it can present the extent of inequality, for example, in income growth, without ambiguity. Moreover, by measuring the growth of GDP per head, the seriousness of overall unbalanced regional development within a nation could be clearly analyzed. Due to the difficulties of indices' selection and weighting, and the multidimensional nature of social welfare aspects, the credibility of a measurement approach with mixed component indices, is vulnerable. Although single indicator measurements might contradict each other, by adopting a few representative indicators, this problem can be effectively resolved, and results would be more reliable in areas such as life expectancy.

Life expectancy at birth is one of the ideal single indices for human development measurement, as it addresses the major concerns of health and quality of life, and as such, has credibility in evaluating overall human development improvement. This is

because the rise or fall of life expectancy at birth will be affected by various factors, such as income, family relationships and health. In other words, the rise of this indicator illustrates personal income growth, harmonious family relations and relatively good health and reflects overall improvements in human welfare. By measuring life expectancy in the different regions, the disparity in spatial human welfare would be clearly demonstrated. In addition, by adopting this approach, the difficult distributional issue, which negatively affects the measurement reliability of both the HDI and GDP, can be avoided.

The Human Development Report 1990 published by the UNDP (1990) has already stressed the important status of life expectancy at birth in studying human development; this is in particular true in the absence of nutritional and health information. Reddy (2007) argues that the attractiveness of the life expectancy approach lies in its capability to measure overall human health standards. Veenhoven's study (1993) suggests that health and longevity have played an important role in measuring human and social development. He states that,

... the better the liveability of a society, the longer the life-expectancy of its members must be. This indicator is certainly quite appropriate where success in providing the biological minimum is concerned. (1993, p. 12)

The importance of life expectancy at birth in analyzing human development is also due to failure of the income approach to measure population health. Indeed, research conducted by Riley (2001) has offered numerous case studies showing the close correlation between high income and low life expectancy. For example, in 1997, in contrast to the average life expectancy 76.7 years in America, the equivalent figures in Jamaica and Belize were only slightly lower, at 74.8 and 74.7 years respectively. However, the income per capita in America (US\$ 29,010) was more than six times higher than Belize (US\$ 4,300) and Jamaica (US\$ 3,440). In addition, in comparison with the average life expectancy of 56 years in Saudi Arabia, the equivalent figure in Sri Lanka was up to 69 years in 1982, even though the income per capita in this country (US\$ 320) was more than 50 times lower than the equivalent figure in Saudi Arabia (US\$ 16,000). Riley (2001) has also cited the evidence of Kerala State, India. Although economically it is one of the poorest states in India, Kerala has achieved the highest average life expectancy in this country.

With regard to China, during the pre-reform period, it recorded very poor economic growth and low income per capita, but average life expectancy at birth rose dramatically. Reddy (2007, p. 50) points out that, "Prior to the adoption of the Open Door policy in 1978, the PRC had achieved extraordinary increases in health indicators, despite its low income." Riley (2001) maintains that the hundreds of 'barefoot' doctors in rural areas have played the key role in contributing to improvement of life expectancy in China. He stresses that the farmers benefited

enormously from these ‘barefoot’ doctors in terms of health care; and their contribution to rural disease prevention and health care, should not be underestimated.

My general view is this: life expectancy is a very important indicator of population health and overall human development. Health has a central position in judging human well-being improvement. Without good health and human longevity, it is very difficult to claim that there has been an improvement in human living conditions and overall human development. In that sense, life expectancy is more desirable than the problematic, mixed component HDI and the over-narrow GDP approach in assessing human development inequality. The life expectancy indicator enables the scholars to estimate overall human development and social progress. More significantly, it can measure whether or not economic growth really contributes to human development. However, the practical difficulty in adopting this approach lies in a lack of reliable data in under-developed regions and countries. As far as this study is concerned, for many economically poor regions within Guangdong, official record of average life expectancy at birth simply have not been kept, let alone updated. At this stage, it is necessary to be practical and realistic in studying regional inequality and an alternative indicator needs to be found. The income approach is a good alternative indicator for measuring human development, being able to largely reflect the improvements in population health and life expectancy. In general, material wealth and economic development are positively associated with the rise of life expectancy. Evidence from many nations has clearly supported this claim. High life expectancy has mainly occurred in the economically rich countries. Wilkinson (1996, p. 41) points out that,

Indeed, it is hard to think of an explanation for rising life expectancy which is not in some way sustained, enabled or supported by economic development.

In fact, even while making strong criticism of the income measurement approach by citing the case studies, Riley (2001) admits that there is some evidence to support the positive correlation between high-income nations and long life expectancy. According to the figures presented in Table 2.5, the close relationship between life expectancy at birth and income per capita is evident. For example, in contrast to the average figure of 66.1 years in the developing countries, the life expectancy at birth in the OECD countries was 78.3 years in 2005. The average GDP per capita in the OECD countries was US\$ 29,197 in 2005, which was more than five times higher than the equivalent figure in the developing countries. In addition, in terms of the least developed countries (US\$ 1,499), the average life expectancy at birth was more than 45 percent lower than the equivalent figure in the high-income OECD nations in 2005 (US\$ 33,831), which was 79.4 years.

Table 2.5: Worldwide Comparison between Human Development and Income Per Capita, 2005

Region	Human development index (HDI) value 2005	Life expectancy at birth (years) 2005	GDP per capita (PPP US\$) 2005
Developing countries	0.691	66.1	5,282
Least developed countries	0.488	54.5	1,499
Sub-Saharan Africa	0.493	49.6	1,998
OECD	0.916	78.3	29,197
High-income OECD	0.947	79.4	33,831
World	0.743	68.1	9,543

UNDP (2008, pp. 232)

Moreover, it is important to highlight that the life expectancy approach certainly is not perfect. The main problem with this approach is that longevity would not be a fair and good indicator in measuring human well-being development in the developed nations, because life expectancy would not change significantly. When life expectancy at birth has reached a certain high level, say, 74 – 78 years, it is very difficult to achieve further substantial rises: even with big improvements in health care and dramatic development of medical technology. For example, as far as the wealthy OCED countries are concerned, the average life expectancy at birth is already notably high; it is difficult to increase this figure significantly and improvements in human well-being would not be reflected clearly in the life expectancy indicator.

Veenhoven (1993) has raised concern about the validity of life expectancy in measuring overall human development disparity. He suggests that life expectancy approach might not reflect the social needs and wants of the individual people among different societies. According to the argument made by Veenhoven, a long life does not necessarily a good and desirable life.

The current research adopts various single economic and human indicators to measure the regional inequality within Guangdong, including income per capita, infant mortality and crude death rate. I believe that adopting several single measurement indicators would be a more reliable measurement of overall spatial disparity than the problematic complex indicator with mixed components. Income growth is the main approach adopted by this study in assessing unbalanced regional development and human development disparity. Ray (1998) highlights the unique advantages of income approach due to its characteristics of simplicity, measurement effectiveness and convenient management. As he points out,

In this sense, the view that economic development is ultimately fuelled by per capita income may be taking things too far, but at least it has the virtue of attempting to reduce a larger set of issues to a smaller set, through the use of economic theory. (1998, p. 9)

Economic development, as demonstrated in GDP growth is still a fair and reliable benchmark of multidimensional development. The overall picture of human and

social inequalities is largely reflected by income disparity due to the positive correlation between these aspects. Therefore, income measurement is still appropriate for analyzing national welfare and regional inequality. Ray (1998) maintains that income per capita is regarded as a benchmark for judging national development performance because many policymakers still see economic development as the fundamental aim of a country.

My view about the HDI and measurement of human development capabilities is that in theoretical terms, the HDI developed by Haq might be a progressive and desirable approach for human inequality measurement. However, in reality, this is simply not the case. Many countries and regions in the world are not ready to adopt this index for regional inequality measurement. This is especially true of the economically poor countries. Some important human development components, which make up the HDI, such as school enrolment rates and life expectancy at birth, are either very difficult to collect or completely undocumented. Without sufficient and reliable data about health, education and other human functionings, sensational claims for promoting the HDI in measuring human disparity are just impractical. In fact, even Sen himself has admitted the challenge of collecting reliable human welfare data in calculating the HDI. Sen (1992, p. 135) states that,

The limits of practical calculations are also set by data restrictions, and this can be particularly hard on the representation of capability sets, as opposed to observed functioning achievements. ... In many situations, practical compromises would have to be made, at least partially.

Indeed, in Guangdong, I have faced many practical difficulties in collecting reliable data on life expectancy at birth for the different regions, especially in the under-developed hill regions. The statistics or health departments in many counties have still not yet collected or kept an official record of life expectancy data. If the data in developed provinces like Guangdong is still not widely available, it is not so difficult to imagine the general challenge in collecting human development data and calculating HDI in other less developed regions of China. Therefore, it is essential to be realistic and we should go back to more sensible and reliable ways for measuring regional and human inequalities within a nation. Only when more data regarding human development and social welfare become available will HDI calculation and analysis be more meaningful.

2.3 Happiness Theory

Happiness Theory¹⁰ is another debatable approach to measuring social and human development inequalities. Richard Layard (2005) argues that happiness is the ultimate

¹⁰ "Happiness is defined as the degree to which an individual judges the overall quality of his life-as-a-whole favourably. Within this concept two 'components' of happiness are distinguished:

and most important goal for a society rather than health, income, political freedom or any other social aspects alone. In terms of his theory, Layard maintains that the absolute priority and crucial role of happiness lies in its self-evident characteristic. Happiness is naturally desirable.

By contrast, he argues that health, income and human freedom are just the intermediary means contributing to the ultimate goal of happiness: termed “instrumental goods” by Layard (2005, p. 113). He seems to argue that these instrumental goals would sometimes contradict each other and need to be kept in balance during the human development process. Layard believes that happiness is the ultimate aim of human development. He claims that these instrumental goals can be inclusively covered in measurements of happiness. If people feel happy, it indicates that they have obtained good health, decent income and enjoyed the benefits of political freedom and social democracy. In general, Layard stresses that happiness is the fundamental criterion by which to measure national welfare and human development disparity. Public health, high economic growth, democracy and other social development aspects, can all be included and judged from the measurement of happiness. Layard proposes seven aspects crucial to happy and satisfying human lives, namely, family relationships, financial situation, work, community and friends, health, personal values and freedom, which he referred to as the “Big Seven” (2005, p. 63).

As far as the credibility of happiness measurement is concerned, Layard claims that happiness is not a subjective and immeasurable approach. Instead, he argues that happiness is objective and can be measured by adopting two methods: questionnaire survey of random population or population brain testing by electrical activity. Thus, the happiness level of the individual person can be obtained and analyzed, and degrees of happiness among different people can be compared. Layard (2005, p. 224) points out,

So happiness is a real, objective phenomenon. Moreover, good feelings drive out bad feelings and vice versa, so that happiness is a single dimension of all waking experience, running from the utmost pain and misery at one extreme to sublime joy and contentment at the other.

Layard criticizes the GDP approach to measurement of social welfare and happiness. During the post-World War II period, the western countries have made huge achievements in economic and industrial development; income per capita has increased dramatically at the same time. However, Layard (2005) found that according to empirical and survey estimation, most people did not become happier in the western developed nations, even though the average income per head doubled and

hedonic level of affect (the degree to which pleasant affect dominates) and contentment (perceived realization of wants). These components represent respectively ‘affective’ and ‘cognitive’ appraisals of life and are seen to figure as subtotals in the overall evaluation of life, called overall happiness.” (Veenhoven, 1993, p. 25)

economic growth was remarkable in contrast to the pre-war period. In general, the people in the western countries have enjoyed better working environments, higher income per capita, better public health care, and many other advancements in personal development. In adopting the income approach, Layard (2005) claims that all countries would be pursuing the goals of increased GDP and economic growth, and the policymakers would regard income growth as the benchmark of national welfare. He believes that the main problem with the economic model lies in neglect of the determining power of social factors in influencing happiness and the wants of people (Layard, 2003). He alleges that,

But in fact the GDP is a hopeless measure of welfare. For since the War that measure has shot up by leaps and bounds, while the happiness of the population has stagnated. (2003, p. 3)

Layard (2005) thinks that happiness should be as important as economic development in judging the success of a country's development policies. He points out the cases of the USA, Britain and Japan to criticize the inadequacy of income approach in estimating human development disparities. Moreover, Easterlin has done significant work on the relation between income and happiness. Easterlin (2001) argues that high income does not necessarily make a person become happier from the angle of life cycle. An upward trend in income would not guarantee the growth of human happiness. Easterlin (2001; 2003) stresses that income growth has failed to continually stimulate the rise of happiness mainly due to the issues of hedonic adaptation and social comparison. Due to natural human characteristics, the material aspirations of human beings will continually develop with the growth of wealth and personal income. In this sense, for the human being, material aspiration and income growth go side by side; material aspiration is never-ending. Easterlin (2001, p. 481) states that,

As a general matter, subjective well-being varies directly with income and inversely with material aspirations. ... Even though rising income means that people can have more goods, the favourable effect of this on welfare is erased by the fact that people want more as they progress through the life cycle.

In addition, people like to compare their own material living standards and personal wealth with those of other people, when they notice the higher living standards and income of others, they do not feel improvements in their own well-being or become much happier. Therefore, income growth has not been able to significantly increase the happiness of the people. The study conducted by Easterlin (2003) has clearly made this point.

However, the research conducted by Veenhoven does not support the arguments made by Layard and Easterlin. Veenhoven (1989) argues that material wealth and economic development are important to human well-being and happiness. Significantly, Hagerty

and Veenhoven (2003) notice that the influence of income growth on the happiness of human beings is getting smaller in the long term, although they highlight the close relation between income growth, economic prosperity and happiness. Drakopoulos (2007) has made a similar argument regarding the positive correlation between income growth, the rise in human happiness and improvement in well-being: maintaining that the strong effect of income growth on happiness becomes smaller when absolute income has reached a certain level. At that time, other factors start to determine the degree of happiness rather than purely material wealth.

To some extent, Layard's argument is reasonable and his Happiness Theory is important. Once again, Happiness argument reveals that the fundamental purpose and goal of economic growth is to enlarge human development potential and human freedom. Above all, people are the overwhelming priority in measuring inequality and judging successful economic development; and the important factors of family relationships and personal values claimed by Layard should be taken seriously as they broaden our knowledge of human development. For example, family relations have a great influence on degree of happiness, particular with regard to children. It is not difficult to understand that children growing up in broken homes would be less happy than other children. Divorce has a negative impact on children's lives. In terms of personal values, Happiness thought suggests that personal achievement in society does account for human happiness. People become happy if their work and personal achievement can be recognized and respected by society. Quite often, it is the recognition of personal values that make a person happy rather than income growth, good health or other social welfare factors. Due to inherited human characteristics and psychological factors, respect for personal values and working achievement is taken very seriously. Moreover, Layard's Happiness approach has added an important indicator for measuring development of political freedom and human rights: personal freedom. Relative to both the GDP and HDI measurement approaches, it is significant progress. In order to fully measure all aspects of human development, political democracy and personal freedom should not be ignored. For example, the extent of happiness for the populations of authoritarian Belarus and democratic Hungary is completely different (Layard, 2005). It is also not difficult to understand that people living in North Korea and Burma may not have experienced much happiness, due to tight controls on political freedom and human development in those countries. When you can freely express yourself and take part in political discussion, you are surely in a happier position than people living without freedom of speech. In addition, Layard's arguments on the value of community and friendship in exploring social development inequality and pursuit of human happiness are also valid.

However, whilst Layard's Happiness Theory measurement is ideal as a theoretical framework, it is not so practical in the real world. One of the key limitations of Happiness Theory is its problematic measurement method. Can happiness be objectively measured as argued by Layard? Is it really possible to analyze the extent of happiness among individual people? Can the seven big factors of the Happiness approach sufficiently measure happiness? The answer to these questions is No. From

my point of view, human development automatically leads to happiness. This causality also runs in the opposite direction as happiness inevitably benefits human development and human freedom. Although Happiness approach is more direct and straightforward than other approaches in measuring human development disparities, the problem with this approach is that happiness is a natural outcome of social welfare improvement and expansion of human potential, rather than an objective indicator which can be empirically measured. Happiness is a complex, comprehensive and subjective feeling involving psychological factors which can not be precisely predicted or objectively measured. Everyone has their own benchmarks for happiness and it can not be uniformly measured by adopting random questionnaire surveys or electrical testing of the human brain as alleged by Layard. Veenhoven (1991, p. 4) states that,

Happiness is insensitive to actual quality of life. Because standards of comparison are arbitrary, the judgements based on them are arbitrary as well. Hence people can be subjectively happy in objectively bad condition, or feel unhappy in good ones. Happiness is a coinage of the brain.

Moreover, Coyne and Boettke (2007) have made strong criticism of the questionnaire survey method of measuring happiness. Their key concern is that the survey method has difficulty capturing all of the sources which trigger happiness. Coyne and Boettke argue that every single person is different and has their own unique characteristics and personality. Therefore, every individual has different subjective feelings and opinions on the concept and meaning of happiness. In addition, they believe that newly emerging opportunities would provide new causes of happiness. In fact, even Layard realizes the subjective characteristic of happiness; he states, "... our happiness depends on our inner self and our philosophy of life." (2005, p. 71-72)

The argument made by Easterlin (2001) is not so appropriate. He argues that the main sources of happiness, such as income, living standards, family and personal health, are very similar among the majority of people. As far as the majority of ordinary people are concerned, income, health and family are probably the three most important sources contributing to happiness. However, it is wrong to ignore potentially unique sources in understanding the happiness of individuals. Sports, personal achievement, charity work, and travel experiences are all very important sources of happiness. In addition, the research done by Oswald (1997) suggests that the impact of unemployment on unhappiness is significant. Special factors affecting happiness should not be underestimated, let alone denied proper discussion. Easterlin has not fully addressed this issue. In fact, in his late study, he admits that,

Thus it seems that there is no iron law of happiness, and no fixed individual setpoint. One's happiness at any given time depends partly on personality and genetics, but also partly on life circumstances. Moreover, the way that happiness changes over the life cycle depends especially on life events. (2005, p. 14)

Moreover, Layard's argument is not so convincing regarding the two methods for happiness measurement. The reliability and usefulness of electrical brain tests in academic research are still debatable and uncertain. We can not totally depend on such problematic data for benchmarks of happiness and human development. As far as the random questionnaire survey is concerned, due to discrepancies on definitions of happiness between individual people, how many questionnaire surveys need to be completed in order to generate a reliable and objective comparison? The exact number might be debatable, but it is certainly not a small number. Under the circumstances of limited time and resources, is this a realistic task? In my opinion, the sources and causing of happiness are impossible to capture fully. Each person has his/her own idea, judgement and understanding about happiness. Everyone has different and unique factors contributing to their happiness or unhappiness. Therefore, happiness is not an objective and sensible approach to measuring both social and human development inequalities. Even Layard (2005) himself admits the unmeasurable and subjective characteristics of happiness.

Moreover, Layard's criticism of the GDP measurement approach is unjustifiable. In theory, his justification of the negative correlation between income and happiness is weak. Layard's limited number of cases in the western developed countries does not warrant his reaching such simple conclusions. The empirical estimation and questionnaire surveys of a few developed western countries are not sufficient to allege the overall negative relation between income growth and happiness. The evidence of the USA, UK, and Japan cited by Layard is not representative and shows bias. In the developed countries, because average income level is high, it is credible to expect that a small income growth would not substantially change the degree of happiness, due to the decreasing affect of wealth and income growth scales¹¹. However, this is not the case for the developing countries, which account for the majority of the world's population, and where slight income change would bring significant improvement to daily living standards. Many people might be lifted out of poverty and change their lives completely. Under these circumstances, income growth naturally leads to happiness enhancement, even if it is accompanied by other factors.

Deng, a farmer interviewee from a northern village in Wuhua, said that his elder daughter could go to secondary school if he could afford the annual 200 *yuan* tuition fee. In addition, his mother could be sent to the people's hospital for an operation if he could afford about 10,000 *yuan* for the operation fee (Interview data 5, October 2006). In Dongshan village, Wuhua, poor teaching environment and a lack of teaching equipment in the local primary school could be significantly improved by about 50,000 *yuan* public investment. There are plenty of similar stories from the poor and

¹¹ In terms of decreasing effect of income growth scale, the study by Layard (2005, p. 51-52) suggests that, "We can examine how far extra income increases happiness for people at different points in the income scale, and we find that the benefit from extra income is indeed less and less the richer the person."

under-developed regions within Guangdong. Income is vital to human and regional development and small amounts of public investment would improve the economically backward environment of many undeveloped regions, whilst slight income growth would dramatically change the lives of many poor people and bring happiness to low-income families. In the study of happiness as a measurement of human development disparity, too much attention has been given to the economically developed nations; while, however, too little research has been conducted in the developing and poor countries. Easterlin (1995) has made a similar claim regarding this issue. In fact, Layard also admits that,

For poorer countries things are different, because people are nearer to the breadline. ... The reason is clear - extra income is really valuable when it lifts people away from sheer physical poverty. (2005, p. 33)

The study conducted by Veenhoven (1989) suggests that between 40 and 50% of Americans already thought they were 'very happy' and 'fairly happy' in 1946 respectively. Therefore, he argues that income growth is unlikely to significantly increase the happiness level due to the natural factor of unhappiness: called "a margin of inevitable unhappiness" by Veenhoven (1989, p. 13). Veenhoven maintains that the effect of income growth on happiness is more prominent in the under-developed countries than the developed ones. He cites the evidence of the rise of happiness in some European countries accompanying economic development during the period from 1948 to 1975.

In fact, some of the case studies of developed countries cited by Layard, are problematic, for example, those in the USA. According to the data given by Easterlin (2001, p. 468), in contrast to merely 16 percent of population being very happy in the income group of less than 10,000 dollars, for the household income group of 75,000 dollars and over, 44 percent of population felt very happy, in 1994. In addition, in terms of the high income group (e.g. 75,000 dollars and over), only 6 percent of the population felt not very happy, rising to 23 percent of population in the low income group (e.g. less than 10,000 dollars). The data suggests that income level is positively associated with happiness. In general, for Americans, a high-level income indicates a happier life. The evidence is contradicted with Layard's argument. In addition, Veenhoven (1989) argues that economic fluctuation has been one of the sources affecting the degree of happiness among American people. Therefore, an insignificant change in happiness does not necessarily imply a negative correlation between income growth and improvement in human well-being.

Oswald (1997) points out that America has experienced a rise of happiness and human well-being improvement accompanying economic development and income growth, even if the rise is fairly small. Moreover, among the nine European countries analyzed by Oswald, excluding Belgium and Britain, up to seven of them had demonstrated a rise of happiness and human development accompanying the economic growth during the period between 1973 and 1990.

My view is this: material wealth and personal income account for happiness; income growth has played an important role in improving overall human well-being. Human living standards can be significantly affected by the actual income level. It is defensible to argue that income and economic development have been the key indicator of overall human development and happiness. Income itself does not contribute to happiness; but is acting as a middleman in the pursuit of overall human development. Without income growth and economic prosperity, it is just hollow to claim happiness and expansion of human development. Coyne and Boettke's study (2007, p. 22) suggests that,

In short, increasing wealth expands the feasibility set of activities available to individuals that allows them to achieve increasing levels of happiness and satisfaction.

The finding of Gallup's new polls (cited in the article 'Where Money Seems to Talk', published by the Economist, July 12th 2007) does not seem to support Layard's theory. The majority of people in the developed nations (e.g. the USA, and Japan) were happier than the people living in the economically poor countries (e.g. the African countries). As far as China is concerned, this country has made impressive achievements in social progress and economic development over the past two decades. The income per capita in China has dramatically increased during the reform period. According to the data presented in Peking University's survey (cited in People's Daily Online, accessed on Feb 24th 2008), the degree of happiness among the Chinese people has risen substantially since 1978. Even the survey study done by the International Gallup Organization demonstrates a similar result. Up to 51 percent of Chinese thought they have become happier and more satisfied with their living standards than in the past. The sample interviewees were 15,000 adults, chosen from provinces throughout China. (People's Daily Online, accessed in Feb 24th 2008)

Moreover, Layard's happiness argument is not totally new. Human development and self-fulfilment would automatically lead to happiness. The personal value of individuals has been emphasized by western society for a long time. Chow's study (1987) highlights this point. Moreover, neither are the so-called 'seven big factors' of happiness measurement proposed by Layard, completely new. In fact, many of these factors have already been covered by both the HDI and GDP measurement approaches. Financial and work situations are directly related to income and economic growth¹²; while health (as measured by life expectancy at birth) is one important component indicator for calculating the HDI. The only new factors in the happiness measurement are those such as family, friends and personal values. Therefore,

¹² DiTella and MacCulloch (2003, cited in Osberg, 2004) have argued the correlation between income level and happiness is positive by studying the sample population residing in the OECD countries during the period between 1975 and 1997. However, by comparing the people living in Chile and Honduras, the findings of Graham and Felton's research (2006) suggests a negative relationship between growth in wealth and rising happiness.

Happiness approach is not really an alternative approach to measurement of human development and regional inequality. Layard (2005) recognizes this limitation. As Layard (2005, p. 113) claims, "... the Big Seven requirements are similar to the personal 'capabilities' that the Nobel laureate Amartya Sen has proposed as the goals of public policy."

In addition to the above drawbacks, the crucial environmental issue and climate change have not been discussed by Layard's Happiness thought and its measurement approach fails to address the impact of environmental issue on human happiness. In conducting international happiness surveys, many companies and research institutions have not paid enough attention to the environmental issue, focusing on income, health, family and personal achievement. Layard seems to believe that the environment would not seriously affect happiness measurement, or is even completely unrelated. However, in my understanding, the environment has great power to influence human happiness, and this issue needs to be taken seriously. Environmental pollution and damage caused by economic and industrial growth would surely make people less happy, or even entirely unhappy. Taking heavily polluted Chinese municipalities as an example, the people living in Taiyuan, Tianjin and Lanzhou will most probably be concerned about their living environment and unhappy with the worsening pollution situation, even though the average income per head for these residents has dramatically increased since 1978. The residents of these three municipalities have suffered from high rates of lung cancer and other respiratory diseases. When you need to take great health risks in your daily life, this would surely affect your happiness. Environmental pollution has undoubtedly affected population health. However, the Happiness thought developed by Layard has not yet taken account of this issue in assessing human development disparity. What a pity!

Layard seems to argue that degree of happiness would not be affected by causal past memories and future anticipation, only long-term or average past memorable events or future human anticipation would account for happiness measurement. The writer is not convinced by his argument. What exactly is meant by "happiness is a feeling and that feeling occur continuously over time throughout our waking life" as claimed by Layard (2005, p. 17)? How can we objectively measure it? It is certainly true that measurement of average past memories and future anticipation is important to achieve happiness. However, this is just part of the story. It is wrong to eliminate the impact of causal or instant events on happiness measurement as they play an important role in causing both happiness and unhappiness. For example, the instant memory of past experiences like relatives' funerals, future anticipation of a wage increase or causal events like the outcome of football match, would all affect the personal feelings of individual people. We can be made happy just by the TV programme we are watching; we might suddenly become unhappy when hearing bad news of a terrorist attack on London or civilian deaths in Iraq. These are neutral human responses. Happiness can happen suddenly. Therefore, it is very difficult to precisely measure all sources and channels leading to happiness and unhappiness. In fact, Layard (2005, p. 17) admits that,

Though our average happiness may be influenced by the pattern of our activities, it is mainly affected by our basic temperament and attitudes and by key features of our life situation - our relationships, our health, and our worries about money.

2.4 Conclusion

The income approach suffers from shortcomings in measuring human development disparities, and addressing the multidimensional aspects of human development and social welfare. Moreover, income growth alone might not guarantee development or strengthening of human potential. Besides these limitations, the issue of environmental damage has not been accounted for in calculation of GDP. The measurement of real GDP growth and per capita income might be not so convincing if the issue of environmental costs has not been properly addressed. However, it is important to note that economic development as mainly reflected in GDP growth is still the most popular indicator for measuring regional inequality and human development disparities as economic development and income growth are the foundation for achieving human development. The income approach is a benchmark for reflecting overall development and inequalities: crossing over from economics to social to political sectors. In reality, the GDP measurement is a good and reliable approach to assessment of human development, which has enabled scholars to explore the issue of regional inequality and understand severe imbalances in regional development. Difference in income growth can account for human welfare disparity and GDP growth has played a crucial role in ensuring human development. Without economic growth, human development is just a vain proposition. The central position of the income approach in measuring full-range development is still valid and justifiable.

In general, both the HDI invented by Haq and the Happiness Theory developed by Layard deserve praise for their contributions to human development theory. Their studies, with their emphasis on the central position of the people, have certainly enlarged our knowledge of the meaning of human development. For example, family relationships and personal values, as proposed by Layard, play an important role in analyzing human development inequality and happiness. Both the HDI and Happiness Theory remind us that the ultimate aim of economic development is to promote social welfare and enlarge human development potential. However, neither the HDI nor Happiness Theory is a real alternative to GDP in measuring overall human and social development. Both these two approaches suffer from various serious drawbacks. On the one hand, many pro-HDI scholars have failed to provide convincing arguments for equal weighting of the three measuring indices. It is widely agreed that the HDI can not capture all aspects of human and social development. As Haq (1999, p. 61) himself stresses, "To conclude, the HDI is neither perfect nor fully developed. It requires continuous analysis and further refinement." On the other hand, as far as the Happiness approach is concerned, it might be desirable in the theoretical measurement

of human development. However, happiness is very difficult to quantify in the real world, due to human diversities and inherent multidimensional aspects. More significantly, Happiness is not a completely new concept. Some of the seven big factors adopted for measuring happiness have been discussed in either the HDI or GDP measurement approach. In addition, the issue of environmental costs has not been addressed entirely by either the HDI or Happiness approach. In order to evaluate full-range development from economic to social and human aspects, the environmental issue must be properly considered.

The single measurement indicator would be more effective in estimating overall spatial inequality than problematic complex indicators with mixed components, such as the HDI. The single indicator could more accurately measure human development disparities. Life expectancy at birth is a good example of this as it can provide careful measurement of human well-being development and population health. As discussed earlier in this chapter, income growth and economic prosperity do not necessarily lead to human well-being improvement; life expectancy is an indicator which can bridge this gap. Degree of regional human disparity would be properly evaluated by adopting the life expectancy approach.

Therefore, neither the problematic HDI nor the over-subjective Happiness approach have been adopted in this study of regional inequality in Guangdong. By contrast, income per capita has been taken as the key indicator of regional economic disparities. The income indicator is not only a fair and popular approach, but also a convenient way to estimate overall spatial inequality. In addition, infant mortality rates and other important human development indicators have been employed in estimations of human development disparity among regions in Guangdong.

Chapter 3 Literature Review Part II: Traditional Economic Geography

3.1 Introduction

The second main hypothesis proposed by this research is to deal with the main cause of regional inequality within Guangdong, including the significance of its geography. Therefore, a literature review with regard to economic geography should be carefully discussed. This chapter offers a critical discussion of the traditional geography theories; it attempts to provide general estimation of their importance to regional development and spatial economic inequality. In precise terms, what are the contributions of geographical location and market proximity to the unbalanced development between the core and peripheral regions? How important is favourable coastal location and market proximity to creation of spatial disparity and unequal economic distribution? These key questions are considered in this study.

The structure of this chapter is organized as follows: Section Two presents the literature review of traditional geographic theories: the specific issues of transportation costs, locational advantages and impacts of new infrastructure have been assessed. Section Three focuses on the critics of the traditional geography. In particular, the Cairncross theory on Death of Distance and Vickerman's infrastructure argument have been highlighted and worldwide evidence is drawn on in analysis of traditional theories. Finally, Section Four forms the conclusion.

3.2 Traditional Economic Geography

3.2.1 Transportation Costs

If these inequalities [**e.g. price and production factor**] are finally abolished by the *ceteris paribus* assumption, only transport costs remain to be minimized. These alone almost always show spatial regularities, and their contribution toward determining location has therefore become the principal item in the ruling theory. (Lösch, 1954, p. 18, bold content added)

Lösch is perhaps one of the earliest renowned contributors to traditional economic geography. Lösch (1954) highlighted the dominance of transportation cost in geography theory. In terms of private enterprises, he pointed out that long distance tends to increase time, transportation and other costs in manufacturing production. Traditional theories suggest that geographical location has a major effect on regional economic growth and spatial industrial distribution via transportation costs. Brakman, et al, (2001) maintain that transportation costs play a central role in economic

geography theory and reflect regional geographical differences. The prospect of minimizing transportation costs tends to attract firms to coastal locations or regions with large markets. Differences of geographical location and transportation costs thus account for varying levels of regional economic development, for example, the sharp economic disparity between the coastal and inland regions within China. As Overman, Redding and Venables (2003, p. 2) state, “distance directly increases transaction costs because of basic shipping costs, the time cost of shipping date sensitive products, the costs of contracting at a distance, and the costs of acquiring information about remote economies.” Beckmann, McGuire and Winsten (1957) point out various operational costs to manufacturers caused by road distance, including those incurred by vehicle operation and maintenance, road tax, petrol, travel time and risk, all of which increase according to distance.

Various locational characteristics have different effects on regional economic growth and lead to the unequal spatial distribution of economic factors, including the market, capital, labour and information. Consequently, geography can influence the regional distribution of industries and other economic activities, and above all, affect economic growth. It seems clear that location of industry close to the sea and large markets is mainly due to the effect of transportation costs. On the one hand, coastal regions generally have good access to the core market with lower transportation cost. On the other hand, in contrast to the peripheral regions with small markets, the core regions¹³ tend to have relatively lower manufacturing costs caused by the effect of economic scale. Low transportation costs cause the manufacturing industries and private firms to cluster in coastal regions with large market potential. Such regions therefore achieve higher economic growth than interior regions with a small market base. Favourable geographical location can be generally summarized in the comparison between the coastal and interior regions, and between the core regions located near to the markets and peripheral regions. This chapter provides more detailed discussion regarding the effects of geographical location (coastal and market proximity) on transportation costs and regional economic growth in the following sections.

¹³ In this chapter, I have utilized the geographical concept – the core and peripheral region – to address the issue of regional economic growth and spatial disparity within a nation. Friedmann (1973) offers a clear definition of both core and peripheral region and their relationship. As he states, “core regions are defined as territorially organized subsystems of society which have a high capacity for generating and absorbing innovative change; peripheral regions are subsystems whose development path is determined chiefly by core region institutions with respect to which they stand in a relation of substantial dependency. Core and periphery together constitute a complete spatial system (1973, p. 67).”

3.2.2 Locational Advantage: Proximity to the Coast or Ocean-Navigable Rivers¹⁴

Relative to hinterlands, coastal regions and regions linked to coasts by ocean navigable waterways are strongly favored in development. (p.184)

Gallup, Sachs and Mellinger's important and influential study (1999) suggests that locations close to the coast or ocean-navigable rivers are more favourable for economic growth; in general, they are more economically developed than their interior and landlocked counterparts. The study offers cause analysis of coastal proximity as an advantage in stimulating economic growth. For the labour-intensive manufacturing and other tertiary industries, the costs of transportation and other production largely depend on geographical location. In order to take advantage of low transportation costs and good access to world markets, manufacturing industries and private companies will naturally be concentrated in coastal areas. Therefore, coastal regions can achieve faster economic growth than those in the interior. Radelet and Sachs's empirical research (1998) also found that shipping costs are significantly influenced by location. Coastal regions with low shipping costs have achieved faster development of foreign trade and economic growth over the past 30 years than landlocked areas. In addition, in the empirical research of Redding and Venables (2001) mathematical calculations of the transportation costs of both intermediate and final goods for enterprises and final profits suggest that the impact of transportation costs on the profits of private firms is remarkable. It was found that a 10 percent increase of transportation costs in inputs and outputs can reduce domestic value-added profits by up to 30 percent; in addition, a 20 percent increase can cause about 60 percent reduction of value-added profits.

Due to the lack of reliable data on transportation costs, the costs of goods shipment has been widely utilized in my research. Shipping costs in specific countries and regions have been demonstrated in order to assess the correlation between transportation costs and economic growth. The research by Radelet and Sachs (1998) has offered empirical estimation of shipping costs in many countries across the world, in particular, in Africa. According to the data presented in Table 3.1 below, in general, relative to the landlocked countries, the shipping costs in coastal countries were far lower during the period from 1965 to 1990. For example, the equivalent figures of CIF/FOB Band were just 2.7, 4.2, 5.0 and 6.0 percent in Canada, France, Ireland and the United Kingdom respectively. However, as far as Chad and Rwanda are

¹⁴ In terms of this research, I have taken the concept given by ESRI and McNally on ocean-navigable river. According to ESRI and Rand McNally (1996 and 1980, cited in Mellinger, Sachs and Gallup, 1999, p. 4), the ocean-navigable rivers can be defined as follows, "... whether a river accommodates vessels with a minimum draft of approximately 3 meters (anything smaller is not considered ocean-going); the point at which a navigable river becomes obstructed by falls, rapids, locks or dams; and whether the river is frozen during winter. The coastline used is free from pack ice throughout the year."

concerned, strikingly, the according figures of CIF/FOB Band were up to 33.6 and 40.6 percent in these two landlocked states.

Table 3.1: The Average CIF/FOB Band in Selected Countries 1965-1990 (%)

Country	CIF/FOB Band	Country	CIF/FOB Band ¹⁵
Algeria	10.0	Myanmar (Burma)	10.1
Australia	10.3	Netherlands	5.6
Austria	4.1	New Zealand	11.5
Bangladesh	11.8	Niger	19.5
Burkina Faso	26.6	Norway	2.7
Cameroon	9.7	Pakistan	9.5
Canada	2.7	Papua New Guinea	13.3
Chad	33.6	Philippines	7.6
Cyprus	10.5	Portugal	10.3
Denmark	4.5	Rwanda	40.6
Finland	4.8	Senegal	13.9
France	4.2	Sierra Leone	12.0
Gambia	16.7	Singapore	6.1
Germany, West	3.0	South Africa	8.3
Ghana	7.8	Spain	6.4
Greece	13.0	Sri Lanka	11.1
Guinea-Bissau	14.8	Sweden	3.5
Guyana	10.1	Switzerland	1.8
Haiti	15.5	Syria	8.5
India	12.1	Tanzania	16.8
Ireland	5.0	Thailand	11.0
Israel	7.6	Togo	19.3
Italy	7.1	Trinidad & Tobago	9.5
Jamaica	15.3	Tunisia	6.7
Japan	9.0	Uganda	10.9
Jordan	12.3	United Kingdom	6.0
Kenya	15.8	United States	4.9
Malawi	33.5	Venezuela	11.3
Malaysia	10.5	Zambia	18.1
Mali	41.7	Zimbabwe	11.2
Mexico	4.8		

(Reproduced from Radelet and Sachs, 1998, pp. 14)

¹⁵ Radelet and Sachs (1998, p. 3) provides a definition of CIF and FOB, "The FOB (free on board) price measures the cost of an imported item at the point of shipment by the exporter, specifically as it is loaded on to a carrier for transport. The CIF (cost-insurance-freight) price measures the costs of the imported item at the point of entry into the importing country, inclusive of related costs, including insurance, handling, and shipping costs, but not including customs charges. The CIF/FOB band, which is our basic measure of shipping costs (SC) is defined as $SC = (CIF/FOB) - 1$."

Shipment costs in Africa have been widely discussed by geographers. For example, in South Africa, due to the high average high transportation costs, the private enterprises engaging in the export-oriented manufacturing and service sectors must pay a high transporting price for imported intermediate inputs, and expect less profit from their exports. These unfavourable conditions tend to discourage foreign investors from investing in South Africa and have had a very negative effect on its economic development (Naudé, 2001). According to Naudé's empirical estimation, in South Africa, during the period between 1988 and 1991, the average CIF-FOB band¹⁶ on imports was about 7 percent; however, the equivalent figures in the developed and developing countries were just 2 and 5 percent respectively. Naudé (2001) points out that the average shipping cost for the developing countries is nearly 50 per cent lower than the equivalent cost figure in South Africa (both the entry and departure shipping costs). Moreover, Radelet and Sachs (1998) point out the sharp difference in shipping costs between coastal and interior African countries. For example, in contrast to coastal Senegal, the manufacturing enterprises located in landlocked Mali, need to pay an additional 48 and 19 percent (US\$ 770 and 770) in order to transport their products to Northern Europe and Japan by water. Relative to the interior Central African Republic, the manufacturing firms in Cameroon can save up to 68 percent (US\$ 1,040) in shipping goods to Northern Europe. In addition, as far as Tanzania and Rwanda are concerned, when comparing the figures for goods shipped to Japan, the shipping cost was US\$ 1,380 in Tanzania; while, the cost in Rwanda reached US\$ 3,880, 1.81 times higher than the equivalent figures in Tanzania (UNCTAD "Review of Maritime Transport 1995", cited in Radelet and Sachs, 1998, p. 15). This is because the labour-intensive manufacturing goods produced in Rwanda first need to be transported to Tanzania's ports by road: resulting in a substantial increase in total shipment costs to Rwandan companies.

Radelet and Sachs found that the equivalent figures of CIF/FOB Band were 11.8 and 17.8 percent between the coastal (80 coastal economies) and landlocked countries (17 interior countries), which suggests the average shipping cost in the coastal economies is around 6 percent lower than landlocked economies. Moreover, in terms of the coastal economies, the average costs of freight and shipping insurance were nearly 50 percent lower than in the landlocked counterparts. Radelet and Sachs (1998) argue that in interior African states, cross-border rail and road transportation, shipping costs, poor local infrastructure, and the high costs of goods insurance all contribute to high overall transport costs. Consequently, these landlocked countries inevitably suffer from slow development of their manufacturing industries and export trade. They tend to be less attractive to foreign investors, and their domestic manufacturing enterprises face strong disadvantages in the international marketplace.

¹⁶ In Naudé's econometric study (2001), the popular CIF-FOB band on import costs has been regarded as the shipping costs.

The data presented in Tables 3.2 and 3.3 below illustrates recent figures of shipping costs in Africa. In terms of the percentage share of freight costs in total value of imports in 2004, relative to the equivalent figure in both developed (3.1 percent) and other developing countries (5.9 percent), the average percentage of the freight costs in value of imports was 9.9, which was more than three times higher than in the developed economies. In the Sub-Saharan countries, the relative share percentage was an even higher 10.3 percent. As far as the African landlocked states are concerned, in Niger and Mali, the percentage share of freight costs in total imported value were up to 23.9 and 24.4 respectively in 2003. In the 2004 cost comparison of product shipment, the freight costs as a percentage of import value in Rwanda and Uganda were 24.1 and 17.4 respectively (see Table 3.3 for more detail). The high average shipping costs in African nations are striking: particularly in the landlocked countries, where general transportation costs might be around seven to eight times higher than the equivalent average figure in developed countries.

Table 3.2: Percentage of Total Freight Costs in Import Value in African Regions, 2004 (In billions of dollars)

Year	Country group	Estimate of freight cost of imports	Value of imports (CIF)	Freight costs as percentage of import value
2004	World total	270.8	9244.7	3.6
	Developed economies	157.7	5928.4	3.1
	Developing economies	75.8	1945.2	5.9
	Of which in:			
	Africa	9.9	151.5	9.9
	Of which:			
	North Africa	4.0	68.7	8.8
	Southern Africa	0.8	13.5	9.0
	West Africa	2.3	32.1	10.9
	Horn and East Africa	1.9	22.9	12.6
	Central Africa	0.9	14.3	9.5
	Sub-Saharan Africa	5.9	82.8	10.3

[Source: Imports based on merchandise imports data from the UNCTAD Handbook of Statistics 2005 (table 1.1); freight and insurance data from the IMF Balance of Payments Statistics on CD-ROM (January 2006); freight ratio estimated as weighted average based on size of economies. This table is not comparable with those found in previous issues of the Review of Maritime Transport owing to changes in sources and methodology.]

Reproduced from United Nations (2006, pp. 123)

Table 3.3: Percentage of Freight Costs in Import Value for Selected African Landlocked Countries in the Selected Years (In millions of dollars)

Year	Country group	Estimate of freight cost of imports	Value of imports (CIF)	Freight costs as percentage of import value
2002	Malawi	101.0	695	14.5
2000	Zambia	108.6	993	10.9

	Western Africa			
2001	Burkina Faso	92.5	656	14.1
2003	Mali	275.8	1130	24.4
2003	Niger	117.3	490	23.9
	Eastern Africa			
2003	Burundi	20.9	157	13.3
2004	Rwanda	61.0	284	24.1
2004	Uganda	288.3	1657	17.4

[Source: Imports based on merchandise imports data from the UNCTAD Handbook of Statistics 2005 (table 1.1); freight and insurance data from the IMF Balance of Payments Statistics on CD-ROM (January 2006). This table is not comparable with those found in previous issues of the Review of Maritime Transport owing to changes in sources and methodology.]

Reproduced from United Nations (2006, pp. 124)

Indeed, according to traditional economic geography, owing to high transportation costs, the under-developed periphery, at long distance from developed core regions, tends not to be favoured by commercial enterprises. As a result, these interior and peripheral regions will show much slower industrial growth and economic development than the coastal regions. As Gallup, Sachs and Mellinger (1999, p. 188) state,

Coastal economies will generally have much lower transport costs than hinterland economies. Countries near core economies (the main providers of capital goods) will generally have lower transport costs than distant countries, so growth is likely to diminish in direct proportion to distance from the core.

Redding and Venables (2001) suggest that remote and landlocked countries face the problem of poor market access to the core regions and high transportation costs of both the intermediate and final goods. Radelet and Sachs's research (1998) suggests that the intermediate and labour-oriented manufacturing industries prefer to locate in regions with low transport costs and good access to the core markets. In manufacturing industries many imported intermediate inputs are required for product assembly. The majority of profits of the intermediate-oriented firms are generated from labour intensive services, such as textile, toy, electronic and computer assembling industries (Radelet and Sachs, 1998). These industries favour coastal countries with good transport access and lower production costs. Radelet and Sachs (1998, p. 2) point out that,

The evidence suggests that high-shipping cost countries will find it more difficult to promote export-led development, even if they reduce tariff rates, remove quantitative restrictions, and follow prudent macroeconomic policies.

Furthermore, Gallup, Sachs and Mellinger's study (1999) has argued that the close correlation between the location of intermediate-oriented manufacturing industries and the coastal regions is due to transportation considerations. The transportation and import of goods account for a significant portion of manufacturing costs and have a big impact on profits. Locating in the coastal areas, the manufacturing industries have the chance to survive and compete in world markets.

Of course, from a political point of view, if inland nations have good relationships with neighbouring coastal countries, to some extent, it is reasonable to claim that the negative effects of topography on the development of trade and exported-oriented industries can be weakened (as illustrated by landlocked Switzerland and its neighbours). Being landlocked might become less important if advantage can be taken of the ports in friendly neighbouring countries, to develop foreign trade with relatively fair transportation costs. However, it is important to note that, whilst good diplomatic relations are useful, they are not able to significantly change the high transportation costs in these nations.

Moreover, Radelet and Sachs's and Gallup, Sachs and Mellinger's arguments are supported by world evidence, as numerous case studies demonstrate the disparity in economic growth between coastal and landlocked countries. Redding and Venables (2001) suggest the strong positive correlation between coastal accessibility and per capita income. They found that coastal regions can achieve per capita income more than 60 percent higher than in landlocked and remote areas. Many large metropolitan and developed areas are situated near to the coast or ocean-navigable rivers. For example, New York, the USA's biggest and most developed region, is also the country's biggest port city. With the exception of countries located in Europe, Gallup, Sachs and Mellinger (1999) argue that none of the world's other landlocked countries are rich in terms of per capita income.

In terms of the African continent, Sub-Saharan Africa in particular, has faced many decades of poverty and slow economic growth. In spite of its past colonial history, corrupt government, and decades of civil war, the key causal factor is probably its geography. There are numerous studies regarding African geography (e.g. Bloom and Sachs, 1998; Limao and Venables, 2001; Sachs, et al, 2004). Most of them strongly argue that the root cause of economic crises, mass starvation and poverty in Africa is remote geographic location, which contributes to the extremely high transportation costs. For example, Malawi and Burundi – two similar landlocked countries – in contrast to the coastal African countries, need to pay an additional 25 and 228 percent in shipping costs for exports to Northern Europe. This is because the goods produced in these two countries first need to be moved either by railway or road before being shipped to international markets from Tanzania's ports (Radelet and Sachs, 1998). Amjadi and Yeats (1995) argue that high freight costs have negatively affected the investment motivation of foreign companies, and made national development goals,

based on resource utilization and industrialization¹⁷, very difficult to achieve. Moreover, Bloom and Sachs (1998) highlight the negative effect of distance from the coast, on transportation costs in Africa. As Bloom and Sachs point out,

In fact, Africa's situation is made dramatically worse by the continent's remarkable disadvantages in transport costs compared to other regions of the world. (1998, p. 20)

In addition, the existence of the Sahara desert has substantially affected overland trade between Sub-Saharan Africa and Europe due to high transportation costs (Sachs, et al, 2004). In India, Gallup, Sachs and Mellinger (1999) point out the massive scale of poverty in the Gangetic Valley; which is at great distance from the coast. Owen (1987) in studies of South Asia, cites the example of Katmandu, Nepal, where the transit distance of goods to the nearest port, in India, is about 500 miles.

In contrast to common belief, economic development in the USA is probably not just due to its efficient bureaucracy, advanced law system and free market economy;

¹⁷ However significantly, rich resource abundance does not necessarily contribute to economic growth. In fact, it might harm economic development. Some scholars argue that state-designated strategies depending on the resource booms to stimulate national economic growth are not as effective as many people believe. Several empirical studies conducted by Sachs and Warner highlight that resource booms have failed to stimulate long-term economic growth. Sachs and Warner (1997) argue that a negative correlation between natural resource abundance and economic growth during the period from 1970 to 1990 was evident. For example, relative to the resource-poor cases like South Korea and Singapore, which have achieved rapid economic development; they argue that many resource-rich nations such as Nigeria and Venezuela have suffered from economic crisis and stagnation. In a later study, Sachs and Warner (1999) offer more countries evidence located in Latin America to support their argument. They point out that although both Bolivia and Ecuador have experienced resource booms, the economic growth rate in these two nations showed a downward trend in terms of GDP per capita. Sachs and Warner (1999) refer to the Dutch disease to address the negative relation between resource booms and economic growth. As they state,

The core of the Dutch disease story is that resource abundance in general or resource booms in particular shift resources away from sectors of the economy that have positive externalities for growth. (1999, p. 48)

Moreover, since the beginning of the 1970s, Sachs and Warner's empirical research (2001) confirms that almost all of the resource-rich nations have experienced economic stagnation. By referring to the case of Botswana, they argue that resource booms in one country have only helped the development of its resource industry; however, the richness of its resource has failed to lead to the growth of other industrial sectors. This is especially true of export-oriented manufacturing industry. Sachs and Warner (2001) stress that natural resource wealth has not contributed to general economic wealth.

geography is surely a factor. In fact, the country has benefited significantly from the geographical advantage of coastal proximity and has achieved rapid economic growth over the past two centuries largely due to its superior location; in precise terms, its proximity to the sea and ocean-navigable rivers. Gallup, Sachs and Mellinger (1999) stress the importance of geographical advantage to understanding American development history. They found that up to 67 percent of the US population live in regions close to either the coast or ocean-navigable rivers (within 100 kilometers). Moreover, Sachs (2005, p. 57) highlights that,

... they [American] forget that they inherited a vast continent rich in natural resources, with great soils and ample rainfall, immense navigable rivers, and thousands of miles of coastline with dozens of natural ports that provide a wonderful foundation for sea-based trade. (Bold content added)

Research into world physical geography done by de Blij and Muller (1993) illustrates the close correlation between the geographical location of large cities in the USA, the shipping industry and coastal regions, for example, Miami, Atlantic City and areas surrounding the Gulf of Mexico. Coastal advantage has given rise to shipping industries, and led to the creation of large cities with high density population. In fact, as we have seen, except for rare cases such as Las Vegas, in America, the majority of the large, wealthy cities with fast economic growth have inherited favoured geography. Glaeser and Kohlhase's research (2004) has made sophisticated analysis of American cities and their geographical location. They found that all of the 20 largest American cities are either by the ocean or are ocean-navigable river ports. This clearly suggests the importance of coastal proximity in regional economic development and spatial distribution of industries and economic activities within the USA.

In an early study by Adam Smith (1910, cited in Radelet and Sachs, 1998), he argues that the Mediterranean basin and its surrounding areas have achieved rapid economic and social development due to coastal accessibility and the resulting prosperous foreign trade. Limao and Venables (2001) claim that the shipping cost of a standardized 40 foot container from a median coastal country is above 50 percent less than from a median landlocked country. In addition, Micco and Pérez's study (2001) suggests that landlocked countries need to pay an additional \$2,170 (nearly 50 percent more) for transporting products. Micco and Pérez also identify that landlocked countries are similar to rising physical distance from the sea by 10,000 kilometers. According to the report published by the World Bank (1998, cited in Limao and Venables, 2001), landlocked and remote countries have difficulty participating in international trade and other global economic activities, mainly due to high transportation cost and poor infrastructure. For such countries, the 1998 World Bank's report found that the average import percentage share in the national GDP was 11; meanwhile, the equivalent figure in the coastal countries was up to 28 percent, in 1995. Between 1965 and 1990, none of the top 15 industrial and other

non-agricultural exporting countries, were landlocked. For example, Pannell and Ma's study highlights that Hong Kong's favourable geographical location should be taken into account for the economic and trade miracle achieved by this tiny island: in particular, the natural harbour with its good accessibility to major world markets. The trade success of Hong Kong can be demonstrated from its leading world position in goods shipment (e.g. the container terminal and its TEUs throughput) (see Table 3.4 below), which in 2003 and 2004, recorded figures of 21.93 and 20.82 millions of TEUs respectively. In 2005, Hong Kong continually remained in one of the top three positions, even if the total millions of 20-foot equivalent containers handled in Hong Kong (22.43) were slightly lower than the equivalent figure in Singapore (23.19). Pannell and Ma (1983, p. 301-02) claim that,

Hong Kong's most significant asset is its location and high degree of accessibility ...

Table 3.4: The World Ranking Top 20 Container Terminals and their Throughput, 2003-2005 (unit: TEUs million and the change of percentage)

Port	Millions of TEUs ¹⁸		
	2005	2004	2003
Singapore	23.19	21.33	18.41
Hong Kong (China)	22.43	21.93	20.82
Shanghai (China)	18.04	14.57	11.37
Shenzhen (China)	16.20	13.65	10.70
Pusan	11.84	11.43	10.37
Gaoxiong, Taiwan (China)	9.47	9.71	8.81
Rotterdam	9.30	8.30	7.10
Hamburg	8.05	7.03	6.14
Dubai	7.62	6.43	5.15
Los Angeles	7.48	7.32	6.61
Long Beach	6.71	5.78	4.66
Antwerp	6.48	6.06	5.44
Qingdao (China)	6.31	5.14	4.24
Port Klang	5.54	5.24	4.80
Ningbo (China)	5.19	4.00	2.77
Tianjin (China)	4.81	3.81	3.01
New York	4.80	4.45	4.04
Guangzhou (China)	4.68	3.31	2.76
Tanjung Pelepas	4.17	4.02	3.50
Laem Chabang	3.81	3.62	3.18
Total top 20	186.12	167.13	143.88

Source: Containerisation International, March, 2006, p. 67.

¹⁸ According to the definition given by the Review of Maritime Transport 2006, TEU is the 20-foot equivalent unit.

With regard to the issue of transportation by air, will good air transport offset the locational disadvantages of the landlocked regions? Although the average price of air transport has been reduced substantially and is continually falling, for the majority of export-oriented manufacturing goods, shipping is still the best and most popular means of international transportation in terms of cost and safety considerations. For this reason, in general, high costs of transportation in landlocked regions will not be changed significantly by the reduction in cost of air transportation.

However, to be fair, it is important to make clear that coastal location does not absolutely guarantee low transportation costs. Infrastructure and inefficient or corrupt bureaucracy can also affect transportation costs in the coastal regions and entirely offset their geographical advantage. In fact, such problems might make the costs of transportation even higher than in the non-coastal countries and seriously affect their trading competitiveness. Radelet and Sachs's research (1998) demonstrates this point. Similarly, Micco and Pérez (2001, p. 20) state that, "the activities required at port level are sometimes crucial for international trade transactions." For example, when comparing highly efficient ports like Hong Kong (China) and Singapore with the inefficient ports of many Latin American and African countries, Micco and Pérez (2001) stress the impact of a port's efficiency on costs of shipping and freight handling. They argue that the low costs in Singapore are due to its effective port management. Moreover, Yeats et al, (1996) argue that some policies (e.g. anti-competitive cargo reservation) implemented by the sub-Saharan African countries have further increased the shipping costs of these countries, using African case studies from Cote d'Ivoire and Senegal.

Of course, due to their port's poor management and backward infrastructure facilities, the location advantage in some coastal countries might be significantly offset and then cause the growth of transportation cost. However, it is entirely inappropriate to turn a blind eye to the determining role played by coastal proximity in lowering transportation costs and influencing the rapid development of trade, industry and national economy. The majority of the world evidence discussed earlier clearly indicates that the regions located near to the coast or ocean-navigable rivers (say, within 100 kilometers) have generally achieved rapid economic growth caused by their low transportation costs which were contributed to by their favourable geographical location and good natural harbours. However, most of the inland countries have failed to realize similarly rapid growth of industries and economy. Natural geography impedes their participation in international and inter-regional trade activities, and limits their potential for economic growth. In addition, issues of bureaucratic inefficiency and poor infrastructure can apply to both under-developed inland regions and coastal areas.

3.2.3 Geographical Advantage: Market Proximity

In a country, for instance, where the great majority of people are too poor to wear leather shoes, setting up a modern shoe factory may be a doubtful business proposition; the market for shoes is too small. (Nurkse, 1971, p. 117)

Nurkse's vivid description suggests that market size¹⁹ is crucial to factory location decisions made by private companies; more significantly, the market largely shapes the economic potential of a region. In terms of one specific region, small size of the intermediate and consumer market can discourage firms from locating there. This is because manufacturing firms which are located in the remote regions with small markets cannot fully generate income and profit, utilize their production capacity or benefit from the knock-on effect of external economies. By contrast, manufacturing enterprises locating in regions near to large markets can entirely eliminate these problems. The industrial producers can serve a large region at low transportation cost; they also can benefit substantially from the increasing returns due to economic scale and the spillover from other externalities, such as technology. The crucial role played by the scale of economies in affecting spatial industrial distribution should not be overlooked. In addition, regions with large markets tend to attract capital, people, intermediate inputs producers and technology, which private companies can then obtain in a short time and at relatively cheap prices. More significantly, firms can benefit from technological spillover and other knowledge externalities generated from regional agglomeration, as research and education institutions generally cluster in the large regions and their satellite areas: for example, Silicon Valley (California) and Zhongguancun (Beijing) are two classic examples. Due to the increasing returns to economic scale generated in the large core regions, Davis and Weinstein's study (1998) highlights the significance of proximity to large markets for private firms and manufacturing industries, calling it the 'home market effect'.

Private companies locating in the regions near to large consumer markets can, moreover, substantially reduce transportation and production costs. Large market proximity also decreases product delivery times and improves the after-sales service to customers. The service enterprises which are close to consumer markets usually have advantages over their rivals, because they can provide better customer service, and can more easily communicate with their consumers (Stafford, 1980). This in turn helps to improve products in order to meet changeable market tastes. Krugman (1991, p. 98) stresses that,

Because of economies of scale, producers have an incentive to concentrate

¹⁹ For purposes of my research, the market size of one region is measured by per capita purchasing power and total local population in that region (the calculation equation is this: the market size = per capita purchasing power * total population). The higher the per capita purchasing power and more population, the bigger the market size.

production of each good or service in a limited number of locations. Because of the costs of transacting across distance, the preferred locations for each individual producer is those where demand is large or supply of inputs is particularly convenient - which in general are the locations chosen by other producers.

Stafford's and Krugman's arguments regarding the importance of market proximity in influencing location decisions are reasonable. The manufacturing enterprises indicate that such decisions do not just depend on reduction of transport costs of raw and intermediate materials. More significantly, lowering the transportation costs of the final products is also essential. The transportation costs of the various inputs and output are the major factor shaping the location decisions of many companies which engage in either manufacturing production or tertiary service industries. It is important to note that transportation costs can be separated into two parts: the raw and intermediate inputs and final goods. The ideal business location is the place where both types of cost can be reduced substantially. Higgins and Savoie (2005) clearly argue this point.

However, it is not an easy task to achieve minimization of transportation costs for both the inputs and outputs at the same time. Therefore, the precise location decision largely depends on the nature of the products and development strategy. The key point is that the firms need to make a clear and balanced location decision, where they can obtain maximum profits and achieve excellent market performance. If their industries involve significant amounts of natural resources and raw materials, as do core mining, iron and steel industries, they generally tend to concentrate in regions with plentiful natural resources. Meanwhile, production made up of the 'ubiquities' feature of intermediate inputs materials, as described by Fenelon (1932), is better located in regions near to the large intermediate inputs and consumer markets.

However, some scholars disagree with the argument about the determining role played by market proximity in shaping firms' location decisions. In fact, for industries such as mining, forestry and farming, Harris (1954) suggests that proximity to regions with raw materials is the most important factor. He cites the case of American industrial distribution to support his argument, and states that,

Certain types of fuel-oriented industries, of course, have been attracted to the centers of energy, either of coal (as in Charleston, West Virginia) or of natural gas (as in Texas), or of water power (as in the Pacific Northwest or the southern Appalachians). (1954, p. 337)

It is important to make the following point. With the development of modern economy and rapid technological innovation, in addition to the continual decrease of transportation costs, the percentage share of raw material utilization in final products, for many industries has been reduced substantially. Raw materials are becoming less important in the final value of products. In contrast to the past, firms tend to use fewer

raw materials in terms of total numbers and quantity; while many industries, such as electronic, textile and toys sectors, have utilized more ubiquitous intermediate-input materials in manufacturing production. My understanding is this: for the majority of firms engaging in either manufacturing or service sectors, proximity to the raw material market becomes less significant than proximity to the intermediate-inputs and consumer markets. Many industries and companies move into regions with large intermediate inputs and consumer markets rather than those with rich natural resources, for example, the food retail industry, financial industries and automobile industry. As Laulajainen and Stafford (1995, p. 59) point out, "Nearness to markets is the single most important factor influencing the location of commercial facilities."

Woodward's case study of Japanese-owned factories (1992) provides support for the argument of the market-based location decision. Woodward claims that Japanese producers tend to choose the regions with large markets. Moreover, an early study made by Lösch (1954), highlights the case of US investment concentration in the border areas between Canada and the USA. The motivation for US companies locating in these areas is their good access to the Canadian intermediate and final goods markets rather than the proximity of raw materials. Being close to the consumer market, according to the argument made by Lösch (1954), US firms not only save delivery time and reduce transport costs, but also, more significantly, US producers can rapidly update market information and better understand their Canadian consumers. By contrast, the regions with huge natural resources might not have as many advantages for economic development as those with large intermediate inputs and consumers markets. Myint (1971, p. 70) stresses that,

... plentiful natural resources are not necessarily associated with prosperity, nor are they a precondition for economic development.

Leamer's case study (1997, cited in Redding and Venables, 2001) and Gallup, Sachs and Mellinger's study of landlocked countries in central European regions (1999), both show that proximity to rich western countries with good transport access plays an absolutely crucial role in explaining their economic prosperity: citing the examples of Poland, Hungary and Czech Republic. In addition, the successful development of East Asia during the post-World War II period demonstrates that proximity to natural resources is not the key factor for economic growth. For example, in contrast to many other Chinese cities, although Shanghai is generally further away from raw agricultural and industrial materials, this disadvantage has not significantly affected its high-speed economic development. More and more foreign and domestic firms choose to locate their headquarters, branches and manufacturing factories in Shanghai and its surrounding metropolitan region. Murphy (1974) supports this argument.

Furthermore, the reason for the firms' concentration in a few core regions is because they have, generally, high population density, and businesses can serve more customers at lower cost. For example, all of the large and developed coastal municipalities within China, e.g. Shanghai, Tianjin and Dalian, have high population

density. The rich and prosperous regions have high density of population. However, regions with high population density do not automatically achieve fast economic development. The Southeast Asian country, Bangladesh is a good example demonstrating the correlation between high population density, under-development and poverty. In terms of Africa, Bloom and Sachs's research (1998) suggests that the demographic factor showing the world's highest youth dependency and fertility rates, is closely linked to slow economic development. Bloom and Sachs argue that high population density will negatively affect the African economy for many years. Gallup, Sachs and Mellinger's study (1999) also suggests that certain regions and countries with high population density tend to have low per capita income and suffer from slow economic development.

3.2.4 Infrastructure, Transportation Costs and Economic Development

Henderson, Shalizi and Venables's study (2001) identifies the important role played by public investment in transportation infrastructure in stimulating local economic development in the under-developed regions. They also suggest that transportation has acted as the key linking means in the process of regional economic development because the manufacturing sector, can only achieve economic development when goods can be freely transited. Moreover, Skinner (1994) argues transportation improvements in remote and peripheral regions can largely overcome the challenge of their inherited geographical isolation and economic backwardness. Consequently, it contributes to balanced regional development between the core and periphery. Owen (1987, p. 12) further points out,

But generally there is a close relation between transportation and economic development because mobility and access are the only means by which the ends in other sectors can be accomplished.

Therefore, construction of good transport links can improve the unfavourable investment environment in remote and landlocked countries, and reinforce competitive advantages, such as low labour cost. Indeed, Bougheas, Demetriades and Morgenroth (1999) argue that differences in regional competitiveness are largely determined by the infrastructure. Martin and Rogers (1995) claim that infrastructure improvement through public investment tends to encourage the private enterprises and industries to relocate into underdeveloped regions. Conversely, if there is failure to provide sufficient infrastructure, regional economic growth will be negatively affected. Kraft, Meyer and Valette (1971) clearly state this point. Therefore, it is argued that infrastructure improvement can also contribute to significant economic growth and living standard improvement of the ordinary people. Redding and Schott's study (2003) indicates that the contribution of infrastructure to lowering transportation cost and shaping regional economic development should not be undervalued. Owen (1987) argues that the negative effect of poor transportation might cause low living standards

and impede people's opportunities to improve their economic condition. As Owen (1987, p. 13) states,

It is equally clear that much of humanity is unable to share in the accumulation of knowledge and technical progress because of the isolation resulting from constraints on travel.

3.2.5 The Effect of Railroad on Economic Growth

In any analysis of the role of transportation in shaping economic growth, railway construction and its associated spillover effects are particularly prominent. The railway has many special and unique advantages over water and road transportation, such as lower transportation cost and greater time saving, Hawke and Higgins (1983, p. 181) state that, "railways were essentially a combination of innovation ... which permitted a marked reduction in the cost of transport." More significantly, the railway has the advantage of being relatively unaffected by the weather and of being safe (O'Brien, 1983). For example, in Russia, both the roads and waterways become difficult to access in winter due to the extremely cold weather (O'Brien, 1977).

It seems that railway construction provides numerous benefits for regional economic growth. The remote and landlocked regions are located far from the coast or navigable rivers; waterway transportation in these regions is simply not possible. In addition, air transportation is very expensive; thus, it is hardly a good option. Although road transportation is relatively cheap and convenient, for many agricultural and industrial products, it just takes too long to deliver the goods to the consumer market. By contrast, unique advantages make the railway the best transportation choice in the remote and interior regions. As Kraft, Meyer and Valette (1971, p. 41) point out,

Railroads are best suited for long-haul transportation and are particularly well suited for carrying bulk commodities. ... These characteristics make railways the best mode for carrying dry bulk commodities in cases where water transportation is not an alternative.

The railway construction can bring dramatically change to a region in terms of the transportation costs reduction, fast industrial and economic development. The great access to the landlocked regions made by the railway has significantly improved their geographical isolation and communication with the other regions. O'Brien's research (1983) highlights the important railway effect in stimulating the economic growth in the landlocked and remote countries.

Railway construction can bring dramatic change to a region in terms of transportation costs and rapid industrial and economic development. Improved rail access to the landlocked regions has significantly improved communication with the other regions. O'Brien's research (1983) highlights the importance of the railway in stimulating

economic growth in landlocked and remote countries. The benefits of railway construction can be reflected in its strong effect on the linkage between the peripheral and core regions. The railway can enlarge the accessibility and utilization potential of natural resources in remote and landlocked regions (Tregear, 1970). Firms located in the periphery are able to transport their goods and raw materials to the core regions with great efficiency, which leads to the reduction of spatial inequality between the rich and poor regions. O'Brien's study (1977) suggests the superior economic advantages brought by railway construction.

The experience of many countries illustrate the important railway effect on regional economic growth. Fenelon (1932) points out that American economic growth, particularly in the West, benefited substantially from railway construction. Similarly, Kraft, Meyer and Valette's study (1971) suggests a strong correlation between railway construction and regional economic growth within America. In the context of Europe, railway construction has not only significantly reduced costs, but also provided greater opportunities for natural resource exploitation in the peripheral regions. O'Brien (1983) identifies the beneficial role played by the railway in regional economic growth within Europe; especially to remote and peripheral places. Moreover, Caron's study of France, Laffut's study of Belgium and Fremdling's study of Germany (cited in O'Brien, 1983), and Metzger's study of Russia (1974) all demonstrate the great contribution of the railway to the economic development of these four European countries. The study by CEDRE (1990, cited in Vickerman, 1991) shows the TGV rail network has boosted the economic development of French regions. O'Brien (1983) suggests that the huge demand and technological requirements generated by railway construction significantly stimulated the development of steel, iron and other supplying industries in South Wales and France, and the development of the engineering industries in Germany, Belgium and Russia. Moreover, Huenemann's research (1984) on the railway's impact in China suggests that railway construction has brought particular benefits to interior regions without water transport access. Detailed discussion by Huenemann of railway and other transportation facilities will be presented in the following chapter.

However, scholars such as Fogel and Hawke argue that the effect of railway construction on the industrial and economic development of a country might not be so significant. Their criticisms will be discussed in the following section. In my view, the railway does not necessarily contribute to rapid economic and industrial growth; the potential economic benefits generated from railway construction should not be overestimated. Meanwhile, to be fair, the railway has contributed to the economic development of many regions. In particular, the remote and landlocked regions enjoy huge economic benefits from the railway effect; the railway reduces the geographical isolation of these regions and provides more opportunities for their commercial links with other regions. Moreover, the railway has enabled these remote regions to lower the transportation costs of both inputs and outputs. Consequently, it stimulates the attractiveness of these underdeveloped regions to the investors, enabling them to obtain more foreign capital and better utilize their natural resources.

The striking correlation between infrastructure improvement and economic growth needs to be taken seriously and infrastructure has, indeed, acted as one of the key stimulating forces contributing to regional economic development. The importance of infrastructure in influencing the economic and social development of the poor regions lies in its power to overcome and cross the roadblock of their inherited extreme geography. Kraft, Meyer and Valette (1971, p. 1) state that,

Transportation systems are designed to overcome the frictions (distance, natural obstacles, etc.) imposed by geography. As such, they shape the distribution of activities and influence the share by which each region contributes to the national product.

Without improving transportation and the other infrastructure facilities, in my view, in the under-developed regions, industrialization might never be achieved. This is one of the few determining conditions for regional industrialization. Not only are the transportation facilities crucial, but also communications, electricity stations, and water. Industrial prosperity must be based on infrastructure development. For example, in the context of Latin America, Wilson (1966) stresses that the strong spread effects from the core to the periphery lie in transportation facilities. Wilson discusses the case of Venezuela. The construction of a new motorway (Tejerías-Valencia Highway) has significantly boosted the spread of industrial development from Caracas into the Maracay-Valencia region. Moreover, the industrial development histories of Europe, America and Japan have clearly demonstrated this point. Therefore, as far as the poor and remote regions are concerned, infrastructure improvement is absolutely crucial for economic and industrial development. Failure to provide sufficient and good infrastructure will inevitably cause slow economic growth and industrial development. Such economic backwardness might never be overcome.

My general argument is that infrastructure improvement is one of the key pre-conditions boosting regional economic growth and substantially reducing spatial disparity between the under-developed (periphery) and developed (core) areas; and without these improvements, rapid economic development is simply too difficult to achieve. For example, in South Korea, the improvement of transportation and other infrastructure via public investment has attracted human talent and capital; it has changed South Korea from an economically poor country with peripheral features to a rich and developed country (Sapir, 1990). Sapir firmly argues that slow economic growth and backwardness in a peripheral country is not destiny and can be changed: infrastructure improvement can make a real difference. Sapir states that,

In the models of Krugman and Venables, the centre and the periphery are defined in terms of location. To my mind, infrastructure and human capital are more important factors in determining whether a country will expand or shrink its manufacturing sector. In other words, being peripheral is not irreversible. (1990, p. 76-77)

Moreover, the early development stories of Shunde, Panyu, Dongguan and Foshan within Guangdong province demonstrate the important contribution of infrastructure to local economic development during the early 'reform and open-door' period. The following Guangdong chapter offers detailed discussion.

Infrastructure has played an important role in affecting the economic development in both developed and under-developed regions. As discussed earlier, if a coastal region has inherited a good harbour without the associated modern port facilities, this region may actually have difficulty in developing international trade and export-oriented manufacturing industries. Infrastructure is in fact fundamentally crucial for both the developed and under-developed regions. The key difference lies in the potential sources for financing these infrastructure projects, in both private and non-private sectors. As far as the economically rich and coastal places are concerned, due to the effect of increasing economy scale, potential advantage of low transportation costs to the international markets, and spillover of other externalities, there is great incentive to invest in infrastructure projects in these regions. In general, the private sector is the key player for financing infrastructure construction in the developed regions, and perhaps the initial source for infrastructure development in the rich regions during their early development stage.

However, to the poor and interior places, due to the effect of small investment returns and poor local financial capability, public investment and subsidies from the state, which are largely non profit-driven, are two crucial sources of finance for local projects of infrastructure improvement. How can the poor regions improve transportation and other infrastructure facilities? The state is the key. This is why the role of a strong state is so important to initially kick-starting economic growth in the under-developed periphery during its early development stage. The role of the state lies in determining capability for infrastructure improvement. When entering the high development stage, the initially under-developed regions might have significantly improved their situation and strengthened sustainable development capability: then the role of the central state in subsidising public infrastructure construction tends to get smaller. At the same time, due to the increasing returns to economic scale and higher investment returns, the private enterprises will start to dominate the local infrastructure development in the under-developed regions. Again, regarding the causality between infrastructure and economic growth, the infrastructure is one of the key stimulating conditions for achieving economic growth: although growth is not automatic: as many other factors are involved.

Many more case studies across the world illustrate the crucial role played by transportation and other infrastructure improvement in shaping regional economic growth and industrial development. In the context of Europe, O'Brien's study (1983) found that the fast economic growth in Europe between 1750 and 1800 was accompanied by rapid improvement of transportation infrastructure. The increasing transportation demand from both industrial producers and farmers was particularly prominent. Moreover, Owen's study (1987) suggests the importance of goods

transportation to American agricultural development. Goods transportation facilities have enabled American farmers to quickly and easily export their agricultural products to both domestic and international markets. Owen further provides a vivid description of good transportation playing an important role in influencing daily food consumption of American consumers. As far as Europe is concerned, the spatial infrastructure inequality at both inter- and intra-country levels is crucial in explaining regional economic convergence within the European Union (Biehl, 1986, cited in Martin and Rogers, 1995). In contrast to America and other developed countries with advanced transportation systems, many developing countries have been facing the problems of backward and insufficient transportation facilities which inevitably lead to the slow development of agricultural and secondary industries. Owen (1987, p. 26) emphasizes that, "without transport at reasonable cost, it is not possible to obtain delivery or meet the cost of seed, pesticides, fuel, and other inputs that modern agriculture requires." For example, the transportation of important agricultural inputs has been impeded by poor infrastructure in Punjab, India.

For decades, India has suffered from backward and insufficient infrastructure, and this has acted as a large roadblock: hindering economic growth and national market integration. Backward infrastructure is not only evident in the remote towns and villages, but also in the large municipalities like Bombay and New Delhi. Owen (1987) points out how infrastructure problems are hindering the economic growth of India, where the majority of the road is un-surfaced. Moreover, India has faced congestion problems in the national expressway system due to insufficient lanes; it is very striking that up to one third of public highways have only one lane. (Apart from the relatively good railway networks), Owen (1987, p. 44) stresses the challenge of poor and insufficient transportation infrastructure in India,

Industrial development in India presents still greater problems. ... the poor roads and the lack of trucks have been major obstacles. ... and the inability of the transport system to meet factory schedules often forces a slowdown of production.

Many African nations offer further evidence of the severe effects of transportation backwardness on local economic and industrial growth. Bloom and Sachs (1998) highlight the extremely poor infrastructure for transportation and communication in Africa. Similarly, Collier and Gunning's study suggests that the negative effects of backward infrastructure lead to slow economic growth and poverty in Africa. Collier and Gunning (1999) argue that the number of telephones per capita in Africa is barely 10 percent of the equivalent figure in Asia; more significantly, infrastructure has also suffered from low quality. In fact, slow infrastructure development has further reinforced the existing high costs of shipping and other kinds of transportation in Africa brought about by landlocked characteristics. Collier and Gunning (1999) also point out the relatively high transportation costs in this continent.

3.3 Traditional Economic Geography: Its Critics

3.3.1 Criticism of Transportation Improvement

Some scholars disagree with the argument regarding the close correlation between the infrastructure improvement and economic growth in the under-developed peripheral regions. By contrast, they believe that the contribution made by transportation construction to stimulate economic growth is not as significant as many people claim. Friedmann (1973, p. 77) argues that,

Governments frequently propose to use transport investments as a development tool for reducing income differences among areas. ... The evidence, however, does not support such a simple generalization.

According to the argument made by Friedmann, both the developed and under-developed regions equally share the benefits of accessibility improvement in the economically poor periphery. Friedmann's study suggests that transportation investment and better accessibility between the core and peripheral regions do not necessarily guarantee the economic development in the periphery. In fact, the spatial inequality between the core and peripheral regions might be even worse. He claims that the transportation improvement in the periphery would be of much more benefit to the core regions than the peripheral regions. This is because the economic assets of the peripheral region such as capital and human talents tend to be attracted to the developed region due to its higher investment returns and scale economies. As a result, the peripheral region faces a further shortage of skilled labour and capital. This is precisely the negative backwash effect on the periphery described by Myrdal. The strong effects of the scale economies and economic agglomeration in the developed regions will continue to contribute to relatively fast economic growth in these places. Therefore, slow economic growth in the periphery tends to persist; under these circumstances, the spatial income disparity between the rich and poor regions will inevitably enlarge. Wilson's research (1966) has also supported this view.

Furthermore, owing to the better accessibility of core regions contributed by the transportation improvement in the periphery, Hermansen's study (1972) suggests that the capital and labour from the periphery will further concentrate in the core regions; as a result, the agglomeration effect enlarges the spatial disparity between the core and peripheral regions rather than the opposite. Some countries' evidence seems to support this argument. For example, the Italian case demonstrates that railway construction contributed to the worsening spatial disparity between the northern and southern regions (Krugman, 1999). The research conducted by Krugman (1999) suggests that the railway caused deindustrialization within the under-developed south, and it reinforced the competitiveness of regional accessibility (e.g. the market and manufacturing inputs) and leading economic position of northern Italy.

Moreover, the economic scale of transportation infrastructure can affect costs and economic growth. It seems to be valid that the density of transportation infrastructure (as measured by the demand for transportation) is more important than infrastructure alone: like high-speed expressways and modern ports. Mori and Koji's research provide a detailed analysis of this issue. They suggest that scale of transportation infrastructure tends to attract industries to cluster in a few core regions, which then lead to industrial agglomeration. As Mori and Koji (2001, p. 21-22) state,

Agglomeration economies are generated by the circular causation between economies of transport density and industrial localization: a greater concentration of industries in a given region generates a larger transport flow through the region, and lowers the cost of transportation via the region due to density economies, which in turn attract a larger number of firms to the region. A successful region eventually emerges as an interregional transport hub as well as an industrial center.

As far as the railway effects on industrial and economic growth are concerned, Fogel's research seems to argue that the role played by the railway in boosting economic growth is actually small. Moreover, in America, he points out that the railway contribution to the manufacturing industries is weak and insignificant. Fogel (1964) highlights the case of America in the nineteenth century. Moreover, Fogel firmly concludes that American manufacturing growth was largely due to technological innovation and market effect rather than railway construction. As far as Britain is concerned, studies of the relationship between British industrial development and the railway effect seem to produce quite contradictory results. On the one hand, the railway construction significantly contributes to the development of industries like iron, steel and timber. For example, Bagwell's study (1974) suggests the importance of railway construction in the growth of British iron and steel industries during the 19th century. Fenelon's case study (1932) demonstrates the strong effect of the railway on the development of British manufacturing, shipping and other heavy industries in the 19th century. Fenelon (1932) stresses that British industrial revolution and expansion could not have been realized without the coal industry development brought by the railway.

On the other hand, Hawke's empirical study of the British railway suggests different conclusions. Hawke claims that the railway's contribution to industrial and economic growth in Britain was actually insignificant. Although he recognizes that the railway's effects on the agricultural industry and passenger transportation were important, Hawke (1970) demonstrates that the social savings of the railway²⁰ seem to be small

²⁰ Regarding the social savings of railway, Hawke (1970, p. 6) points out that, "Although not explicitly taken from the general corpus of economic theory, this is an *ex post* analogue to the *ex ante* concept used in cost/benefit analysis, and it is measured as the difference between the actual cost of the transportation services of a given year provided by the railways and the hypothetical cost of those same services in the absence of the railways."

in contrast to the total national income. Moreover, Hawke's study plays down the importance of railway construction in the development of British iron and steel industries: claiming that the railway's contribution was, in fact, minor.

One of the strongest critics of the pro-transportation scholars is perhaps Vickerman. His study illustrates the negative effects of transportation improvement on regional economic growth. Vickerman (1991) alleges that there is no hard evidence to support the argument for the influence of transportation improvement in stimulating regional economic convergence. Moreover, Vickerman, et al, (1995) disagree with the argument that infrastructure improvement will automatically bring economic prosperity to the under-developed peripheral regions. In his later study, Vickerman (1998) further suggests that the strong effect of the increasing economic returns generated in the core region might offset the economic opportunities of the peripheral regions brought by transportation improvement. Vickerman's study is supported by Fujita and Mori. Their research has also pointed out the negative infrastructure effects on the economic growth of the less developed periphery. Fujita and Mori (1996) claim that the inter-regional infrastructure improvement between the core and periphery will reinforce the competitive advantages of the industries located in the core region; it may attract more industries to the core region from the periphery.

In Europe, regional income disparities had been rising substantially between 1970 and 1988. Vickerman (1991) demonstrates evidence of a negative correlation between public infrastructure investment, economic growth and regional convergence. His study suggests that the role played by public infrastructure investment in encouraging the process of regional development within the poor region might be weak. Moreover, regional disparity might even increase with improvement of transportation and other infrastructure in the least developed regions. Vickerman (1991, p. 1) suggests that,

However, long periods of regional development policies based on the creation of infrastructure in lagging regions have failed to make much impression on regional divergence. ...This might suggest that infrastructure is irrelevant to the real determinants of regional growth and development.

From my point of view, Vickerman's study has provided some useful information on regional economics. It is certainly true that infrastructure is not everything and that infrastructure improvements alone cannot guarantee the economic development in the under-developed periphery. Improvements which link the developed and under-developed regions not only provide better access and greater economic opportunities for the peripheral regions; but also more significantly, offer good transport access for the developed regions to enter the peripheral regions. It encourages more private companies to agglomerate in the developed regions, and allows them to better exploit natural resources and expand sales markets in the under-developed periphery. Therefore, the leading economic position of the

developed regions tends to be reinforced: and as a result, the spatial disparity between the core and periphery might, in reality, be further increased.

However, significantly, Vickerman's theory is not above criticism: in fact, it is problematic. Vickerman is too pessimistic about the contribution made by transportation and other infrastructure facilities to overcoming economic backwardness in the under-developed periphery. Regional evidence regarding the significance of transportation improvement on economic growth within Guangdong does not support Vickerman's argument. The detailed case studies of Guangdong are presented in the following chapter. In my opinion, Vickerman seems not to properly realize that the infrastructure facilities are a key factor in encouraging commercial communication between the poor and rich regions; infrastructure improvement is able to substantially change the geographical isolation in the under-developed periphery. O'Connor (1971, p. 195) argues that,

It is widely accepted that transport facilities are among the most important factors affecting the distribution of economic activities ...

Although favoured geographical location with low transportation cost and good infrastructure are not the only pre-conditions for achieving regional economic development; without question, they are crucial factors in shaping regional economic growth and spatial industrial distribution. Based on theoretical discussion and worldwide evidence demonstration, one thing becomes clear: in any level of the geographical space – whether international or inter-regional, relative to the coastal regions, the landlocked and interior regions generally have been facing much higher transportation costs and poorer access to the core regions due to both distance and insufficient infrastructure facilities. Radelet and Sachs (1998) rightly point out this issue.

As a result, when considering the remote landlocked regions, transportation improvement might be their only real hope of overcoming geographical isolation and taking part in both the inter-regional and international commerce and trade. Without such improvement, it is very difficult (if not totally impossible) to achieve fast economic growth and improve living standards for the ordinary people in these regions. The evidences from worldwide sources discussed earlier illustrates this point. Vickerman's case study is too general and over simplified; more significantly, he mainly focuses on the European continent. The European case alone cannot tell the whole story of infrastructure improvement and regional development. In Europe, the transportation and other basic infrastructure facilities are generally advanced and sufficient; therefore, the positive effect generated from transportation improvement in stimulating economic growth might not be as significant as in many other poor Third World countries. Under the circumstances, Europe is hardly a good example to demonstrate the importance of transportation improvement in influencing regional economic growth and reducing spatial inequality. The limited case studies presented in his research do not justify Vickerman's simplistic conclusions. The correlation

between transportation improvement and economic growth is far more complicated than he suggests.

3.3.2 Cairncross's Theory: The Death of Distance

Moreover, owing to the reduction or 'disappearance' of distance in the future, some scholars (e.g. Friedmann, 1966) argue that the issue of transportation costs would become less important in firms' location decisions. They seem to suggest that spatial industrial distribution, economic growth and regional inequality will not be seriously affected by geographical location or transportation costs in the near future.

The general argument proposed by this sort of theory is as following. Scientific development and technological revolution will not only dramatically reduce the costs of transportation and communication, and change the way of economic development and manufacturing production; but also more important, the way of human communication and living. Glaeser and Kohlhase's research highlights the continual decrease of goods transportation costs during the twentieth century (up by 90 percent) and the weaker role played by transportation infrastructure in regional growth due to technological change and scientific revolution. In their research, by eliminating the crucial role of transportation infrastructure (e.g. the rail and water transports), Glaeser and Kohlhase (2004) stress that technology innovation would be the major contributor to balanced regional development; above all, technology can reduce the spatial economic disparity between the urban (core) and rural regions (periphery).

Frances Cairncross is one of the main proponents of distance disappearance theory. In her book, *'The Death of Distance'*, Cairncross (1998) firmly argues that, in the near future, geography will no longer be taken into account, and it will become less important for regional economic growth; she refers to it as the 'the death of distance'. Cairncross claims that the death of distance can offer commercial companies more incentive to relocate to the periphery rather than the core region. Cairncross (1998, p. xi) states that, "No longer will location be the key to most business decisions."

According to the Cairncross model, technological innovation and change is the key factor in ending distance dominance in the firms' location decisions and the economic growth of a country. The fundamental technology revolution, with the development of Internet and wide utilization of television and telephone, will completely change the ways of industrial production, transportation and customer service provision. The new technology will also reshape the unequal spatial distribution of economic activities and lead to regional income convergence. Cairncross (1998, p. 1) points out that,

Technological change has the power to revolutionize the way people live, and this one will be no exception. It will alter, in ways that are only dimly imaginable, decisions about where people work and what kind of work they do, concepts of national borders and sovereignty, and patterns of

international trade.

Cairncross further provides vivid and comprehensive discussion of the dramatic impact of new communication innovation on the operations of the private company, product distribution and customer service provision. In my view, Cairncross is only partly right. Of course under certain circumstances, the technological revolution and new science of telecommunications substantially reduce the costs of transportation and other manufacturing processes. Moreover, the technological innovation has significantly changed methods of industrial production and business operation. Many kinds of products and intangible services can be delivered electronically via the telephone and Internet rather than in tangible ways: by road, rail or air. The scientific revolution has contributed to the development of new industries, such as Internet banking, on-line and telephone customer service (e.g. the call centre). Furthermore, the revolution of science and technology has enhanced industrial productivity and changed ways of human communication.

Nevertheless, some parts of the Cairncross theory are unconvincing and problematic. There are certain limitations to her model of the Death of Distance. The key questions regarding her theory are these: will technological change lead to the tendency of spatial equality and regional convergence? Is it justified to argue that distance will be completely eliminated from regional study in the coming future? In addition, can the innovations of technology and communications really cause the death of distance? I think not. Without means of transportation, many kinds of business and trade cannot completely function; numerous tangible products and services cannot be delivered to the customers. The face-to-face method of customer service is still essential and required by many tertiary industries. As Laulajainen and Stafford's study (1995, p. 62-63) highlights, "The friction of distance is always present. It costs money, time and effort to get from one place to another."

For example, it is hard to image how the automobile maintenance service can be completed without guidance from professional repair people. It is hard to believe how Christmas gift toys (e.g. the baby bear), most of which are produced in China, can be sent to children in Europe and North America without physical delivery by road, sea and air. Therefore, the involvement of distance and transportation costs for the service, goods and people are simply inevitable. In fact, Cairncross herself recognizes the problem of her thought. Some of the products and services cannot be delivered without physical distribution and transportation. Cairncross (1998, p. 213-14) admits that, "But where a product requires physical distribution, the impact of the death of distance will be lessened."

The power of technological innovation might have significantly changed business communication for the service industries. However, for the manufacturing and other labour-intensive industries, this is hardly true. As far as both manufacturing and service industries are concerned, it is justified to argue that their communication costs have been reduced substantially by the new technology effect, but it is not necessarily

associated with cost reduction of physical transportation of goods. In fact, product transport costs still persist and heavily account for the spending of the majority of industries. Therefore, to the manufacturing enterprises, transportation costs and geographical distance will continue to be important for location decisions in the future. Henderson, Shalizi and Venables (2001) criticize the thought of the Death of Distance. Similarly, Venables's study (2001) strongly doubts the effects of technology and communications innovation to achieve such an effect. Venables argues that transportation costs and distance will not disappear completely in the future. In fact, the opposite is true: distance and transportation costs will continue to play an important role in influencing the location decisions made by the companies and continue to affect regional economic growth. As Venables (2001, p. 2) claims,

This view of the effects of ICT [**Information and Communications Technologies**] is misleading ... First, new technologies will have a mixed and complex effect on the costs of distance. Some activities can be digitized and supplied from a distance, but most cannot. (Bold content added)

Both these studies should be taken seriously: their arguments are plausible. Their research indicates that distance will continue to play a crucial role in shaping the production location decisions of private businesses; and affecting spatial industrial distribution via transportation costs. Moreover, distance will still influence regional economic growth and spatial inequality. Venables thinks (2001) that for many poor countries, the root of backwardness is their unfavourable geography and their long distance from the core regions. Instead of narrowing spatial income inequality, he argues that technological progress has contributed to regional industrial convergence and widening spatial inequality. For example, Venables argues that the nineteenth century transportation revolution caused the concentration of industries and other economic activities. In addition, with the rapid development of modern technology, the financial industries seem to be more and more relocated into a few core places (Venables, 2001), such as London, Tokyo, and New York.

My general argument is that technological change and communication innovation are not so powerful and influential as Cairncross claims in causing the death of distance and reshaping spatial industrial distribution. Cairncross overestimates the technological innovation effect: and her theory seems not to be supported by world evidence to any appreciable extent. For example, the case studies of other developing countries made by Radelet and Sachs (1999, cited in Mellinger, Sachs and Gallup, 1999) demonstrate that the geographical advantage of American coastal regions was not affected by the technology innovation and change. Coastal regions still dominate national economic and export growth. Moreover, Venables (2001, p. 16) stresses that,

New technologies allow dispersion of activities that only require 'conversational' transactions, but might also increase the complexity of production and design process, and hence increase the proportion of

activities that require 'handshake' communications.

It almost seems to be the case that Cairncross has not used her model to test the current world situation. Actual case studies of individual countries suggested by Cairncross are very rare. Based on a few industries like the Internet, telecommunications and advertising, she alleges that the technological innovation will fundamentally change the method of firms' production and substantially reduce transportation costs. However, Cairncross offers very little evidence to support the idea that the technological revolution has actually encouraged rapid economic development in the poor countries and brought distance analysis to an end. It is certainly the case that such innovation has dramatically changed industrial production and enhanced manufacturing productivity during the period post-World War II; it has also contributed to the relocation of some industries to poor and remote regions. However, such redistribution into the under-developed regions did not necessarily lead to fast economic growth, or reduce the disparity between the developed and least developed regions. The theory of the Death of Distance cannot be properly supported without comprehensive world evidence to back it up. In fact, the technological revolution tends to further contribute to economic growth in the core regions and reinforce their economic competitiveness. As a result, this theory developed by Cairncross is weak and unreliable; her industrial case studies seem to be just anecdotal. Moreover, world evidence demonstrates that the roles of the distance and transportation costs are still crucial in affecting spatial industrial distribution and regional economic development. Venables (2001) has made a fair judgment on the credibility of distance death thought. Although the contribution made by new technologies to economic growth is significant, he stresses that it will not cause the disappearance of distance or changes to the fundamental principles of economic geography.

Therefore, the improvement of transportation and other infrastructure facilities are still essential as a shortage of transportation facilities can seriously hinder regional economic growth. Owen (1987) states that the operation and practice of all economic sectors in the economic development process rely on transport: he refers to the importance of transportation as the 'catalytic role'. Without tackling the accessibility problem, Owen (1987) stresses that nothing else can be done to seriously deal with the challenging issues of poverty and inequality. Similarly, Isard's study (1956) also recognizes the important role of transportation in industrial production; and the dependence of economic activity on transportation facilities.

In conclusion, world evidence suggests that transportation improvement will dramatically overcome the challenge of high transportation costs and long travel time between the peripheral and core regions and will substantially increase mutual communication. Moreover, it will enhance the firms' motivation to relocate into the remote and poor regions. Consequently, it will bring the visual and real hope for the under-developed periphery, and enable these regions to overcome the causal

correlation between slow economic growth and landlocked geography. Wilson (1966) has made proper conclusions on the importance of transportation.

3.3.3 Criticism of the Location Advantage

In his *'In Search of Prosperity'*, Dani Rodrik (2003) firmly argues that state policies, rather than geography, have played the determining role in contributing to regional economic growth and spatial disparity reduction. In precise terms, the rule of law and guaranteed property rights are the key; their crucial positions in affecting economic development should be fully appreciated. Rodrik (2003, p. 8) gives it the "utmost importance in initiating and sustaining economic growth". In particular, he raises the case studies of Botswana, China, Australia and Mauritius to support his pro-institution argument on regional economic development. Rodrik (2003, p. 10) suggests that,

Institutions that provide dependable property rights, manage conflict, maintain law and order, and align economic incentives with social costs and benefits are the foundation of long-term growth.

Botswana is perhaps the strongest case proposed by Rodrik in supporting his pro-institution arguments. According to the arguments made by the pro-institution scholars, Botswana is a landlocked African country with non-favourable geography. If traditional geography theory is correct, relative to coastal countries, Botswana should be economically poor. However, it has achieved fast economic growth (7.7 percent annual growth rate) and relatively high PPP-adjusted per capita GDP (US\$ 5,796) during the past 35 years (Acemoglu, Johnson, and Robinson, 2003). The study by these writers suggests that the state-oriented development policies have been crucial in contributing to the rapid economic development in Botswana. In addition, in this nation, the strong role of institutions can be reflected from the huge investment in public goods and services, such as infrastructure construction, basic education and the public health system.

As far as Botswana is concerned, it is clear that the country has made impressive economic achievements during the past three decades. However, to what extent, can the high-speed economic performance of Botswana be attributed to the state? Is it really justified to argue that the economic development in Botswana should be entirely credited to its government, as claimed by Rodrik? In my view, his argument is not very convincing. In fact, I am very concerned about the real governing capability of the Botswana government, and citing the case of Botswana as evidence for the role of a strong state in economic development is very controversial. If its government is really making a great contribution to Botswana's economic growth and social development, why is the AIDS/HIV disease so rampant and why do so many people still suffer from this disease in Botswana, if the government has invested so heavily in public health? (Apparently, this is a good example of government failure!)

Sheila Tlou, the Minister of Health of the Republic of Botswana, claims that Botswana is the country most badly affected by the HIV/AIDS epidemic in the world. Today, she stresses that the majority of people of Botswana have at least one friend or family member who has lost her/his life to this deadly epidemic (cited in 2005 Botswana Second Generation HIV/AIDS Surveillance: Technical Report²¹, 2005). This is perhaps the most detailed report regarding HIV/AIDS in Botswana, and has clearly demonstrated its prevalence in this nation:

HIV/AIDS in Botswana has become a serious development and social problem impacting negatively on all sectors. The President, H. E. Festus Mogae has described the HIV/AIDS epidemic a national emergency, and a war which the government and people of Botswana are determined to win. (2005, p. 18)

In Botswana, in addition to HIV/AIDS, high unemployment rate and serious regional economic disparity are all apparent, and Botswana's government has completely failed to deal with these problems. Acemoglu, Johnson, and Robinson (2003) state that it is uncertain whether or not the government in Botswana is capable of addressing the severe challenges of AIDS, rising regional disparity and high unemployment rate.

In fact, Botswana's rapid economic growth has significantly benefited from its huge natural resources, and associated international trade. Acemoglu, Johnson, and Robinson (2003) state that nearly 40 percent of Botswana's industrial output has been generated from the diamond trade. Moreover, according to official government source (cited in the website of the Republic of Botswana, accessed on 15th November 2007), the diamond trade is the key source of the national income of Botswana. The revenue generated from the diamond industry accounts for the majority of total exports (77 percent) and almost half of total GDP of Botswana (45 percent).

Without doing any specific empirical statistical analysis, more than 40 percent of total GDP lies in the diamond mining industry, and the rapid economic development in Botswana largely depends on one natural resource. More significantly, because of the diamond trade, Botswana's government can develop sufficient financial capability to invest in public services like health and education. It is reasonable to argue that the example of Botswana's impressive economic performance is exceptional, or maybe just good luck in having such a valuable natural resource.

Furthermore, it is important to note that geography has the determining power in shaping the precise location and distribution and quality of natural resources; and high-speed economic growth in Botswana can largely be attributed to the key issue of

²¹ The Technical Report '2005 Botswana Second Generation HIV/AIDS Surveillance', is published by the Department of HIV/AIDS Prevention and Care, Ministry of Health of the Republic of Botswana with technical help and cooperation from the United Nations and World Health Organization.

its geography. Although Botswana is landlocked, the rich natural reserves might entirely offset this other geographical disadvantage. Rodrik (2003) points out the geographical significance of natural resources.

My general view is that the story of Botswana's economic success is basically a case of diamond exploitation and trade; the fundamental cause of economic development in this country is a rich natural resource rather than the role of a strong state. Rodrik, Acemoglu and other scholars seem to be too optimistic about the governing capability of the Botswana government. Apparently, they have overestimated the role of state in promoting the economic growth of Botswana. In fact, even if he firmly argues the determining position of the institutions in stimulating foreign trade, industrial development and economic growth, Rodrik (2003) has at least recognized the importance of geography in causing spatial income inequality, and the influence of geography on government policies. Countries which are landlocked or have other unfavourable geographical characteristics have suffered from slow economic growth. The majority of these countries have remained in an economically backward position. As Rodrik (2003, p. 5-7) himself admits,

Geography is an important determinant of the extent to which a country can become integrated with world markets, regardless of the country's own trade policies. A distant, landlocked country faces greater costs of integration. Similarly, geography shapes institutions in a number of ways.

However, my research does not suggest that the state and institutional policies have nothing to do with regional economic growth: or that the state should not be taken seriously or even omitted from regional economic study. From my point of view, the state-oriented policies have a lot of power in shaping balanced regional development; in particular in the task of boosting infrastructure construction. As discussed earlier, public investment from the central state in improving the transportation facilities within the under-developed regions is absolutely crucial. More significantly, in terms of the under-developed periphery, it seems to be the case that financial support from the government is the only possible and reliable way to boost the development of transportation and other infrastructure facilities.

The literature review framework of this study emphasizes critical discussion of the geography effect on economic growth and regional disparity. Therefore, it is not my intention to further discuss the role of the state in my research. In the following chapter, based on the regression model and other empirical estimation, the general role of the state in influencing regional economic growth in Guangdong will be addressed.

3.4 Conclusion

In the literature framework presented in this chapter, I analyzed the significance of traditional economic geography in assessing the regional economic growth and spatial inequality within a nation: citing evidence from throughout the world. There are some proper conclusions to be drawn from this chapter. On the one hand, coastal location and transportation costs have played an important role in influencing economic development. In general, the regions which have achieved rapid economic growth and high income per capita benefit substantially from good access and low transportation costs caused by geographical advantages of coastal or market proximity. Without such advantages, these regions would have difficulty in achieving such high-speed economic growth and industrial development. As Gallup, Sachs and Mellinger (1999, p. 182) stress,

We believe that geography - along with economic and political institutions - continues to matter for economic development. Thus, geographic considerations should be taken into account in econometric and theoretical studies of cross-country economic growth, which have almost completely neglected geographic themes.

In general, relative to the landlocked counterparts, the coastal regions have good access to core markets and lower transportation costs which provide the private firms with greater incentives to agglomerate in these areas. The majority of world evidence shows that the regions located near to the coast, due to favourable geography, have generally achieved fast economic growth, which, however, most of the interior countries are unable to realize.

Furthermore, in the under-developed peripheral regions, infrastructure construction plays an indispensable and crucial role in speeding up economic growth and enhancing living standards of the ordinary people. The importance of infrastructure in the under-developed regions lies in its great power to break the natural geography roadblock in these regions. The striking correlation between infrastructure improvement and economic growth needs to be addressed: although, locational advantage alone cannot automatically lead to regional economic growth. Neither can infrastructure improvement guarantee economic development in the under-developed periphery, and spatial disparity between the core and peripheral regions might in fact increase.

The literature review of traditional geographical theories reflects that geographical location and transportation costs are important to economic development. The majority of world evidence suggests that the regions which have achieved rapid economic growth and high income per capita benefit substantially from their locational advantages. Therefore, it is logical to expect that the coastal regions within Guangdong, which are a short distance from large markets, would have achieved

rapid economic growth in comparison to its inland and mountainous regions. In seeking to test the second hypothesis, a theoretical framework of traditional economic geography is applied to the context of Guangdong in the following chapter.

Chapter 4 Literature Review Part III: Regional Economic Divergence

4.1 Introduction

In studying the regional inequality within Guangdong during the reform period, it is logical to investigate the development tendency of this regional disparity. Will it demonstrate a divergent or a convergent trend? Therefore, literature regarding regional divergence will be critically discussed in this chapter. There are two main sections included in the chapter, the first of which focuses on the analysis of new economic geography. What is the role of historical incidents in influencing regional convergence and spatial economic inequality? How much does economic agglomeration affect the unequal industrial distribution among the regions? How does industrial cluster influence the core-periphery pattern of regional economic disparity? These important questions will be investigated in the first section.

Section Two presents general discussion of agglomeration effect and regional economic growth. First, I examine the effect of regional convergence on economic growth and spatial inequality within a country. Secondly, the analysis between history and regional convergence is presented. Thirdly, this section provides correlation analysis between economic factor concentration, agglomeration and spatial disparity. The potential impacts of technology diffusion, the backward and forward linkages on shaping regional convergence have been emphasized. Fourth, I examine the challenges of the main opposing forces to regional agglomeration. Fifth, the self-reinforcing pattern of agglomeration effect and the contribution of path dependence theory to the new economic geography are assessed. In addition, the main criticisms of new economic geography will be analyzed in this section.

The second section also highlights the other key theories on regional income divergence. Myrdal's theory of spread and backwash effects, Friedmann's regional economic polarization and Hirschman's growth pole thought are discussed. Moreover, general introduction of endogenous growth theory is also included here. General conclusions are drawn in the final section of the chapter.

4.2 New Economic Geography

4.2.1 The Agglomeration Effect and Economic Growth

As discussed in the previous chapter, although the traditional theories of geographical location and transportation costs have provided a good explanation of the causes of regional economic growth; they are unable to give satisfactory answers to the

questions of spatial industrial clusters and agglomerative economies. Why do regions with similar natural resources, transport costs and other local features experience very different paces of economic growth? In addition, the traditional geographical theory has failed to explain why regions with similar geographical conditions develop dramatically different industrial structures and demonstrate unequal wealth distribution. Ottaviano and Puga (1997, p. 707) argue that,

The determinants of spatial differences in the patterns of production have traditionally been presented in terms of endowments, technologies, or policy regimes. Such explanations, while relevant, fail to explain why even a priori similar regions can develop very different production structures.

For example, as far as China is concerned, Wuhan municipality has a similarly superior location to Shanghai, which is also located on the Yangtze River. Since ancient times, Wuhan has become an important transportation centre for the surrounding regions of the Yangtze River, served by road, railway and water. The advantage of location has significantly reduced the transportation costs of Wuhan; which allowed it to become the regional commercial centre. However, in contrast to Shanghai during the reform period, why has Wuhan failed to achieve rapid economic growth and regional concentration of economic activities? When comparing their similar coastal location, low transportation costs and inherited commercial history, why has Shanghai become the economic, trade and commercial centre of China, rather than Wuhan? Zhang, Xiong and Shen (2002) state that Wuhan has been a crucial regional centre and sea port in Hubei since the East Han dynasty (*Donghan*). They further add that,

Wuhan, located in the middle of Yangtze River, is the key transportation network for the east - west and south - north regions within China. Two famous rivers - Yangtze and Han [**Hanshui**] cross in Wuhan. (2002, p. 5-6, bold content added)

Analyzing regional agglomeration without understanding its meaning is inappropriate. It is important to address the definition of agglomeration before further discussion of the new economic geography is undertaken. Malmberg and Maskell (1997) argue that agglomeration is due to the industrial linkages and economic concentration which can dramatically reduce business transaction costs and improve profit performance for the private enterprises. Thus, private agents have gained sufficient incentive to locate close to other firms: which naturally leads to economic agglomeration. Furthermore, Ottaviano and Thisse (2002, p. 18) defines agglomeration as, "the outcome of cumulative processes involving both supply and demand". The most clear definition and conceptual framework of agglomeration is probably proposed by Parr (2002). Parr suggests that the rise of economic agglomeration is closely linked to cost savings and profit maximization of the private agents. In addition, the increasing returns to

economic scale and various economic externalities are directly correlated with the development of regional economic agglomeration. As Parr (2002, p. 718) maintains,

Agglomeration economies are regarded here as cost savings to the firm which result from the concentration of production at a given location, either on the part of the individual firm or by firms in general.

The so-called 'new economic geography' (NEG) attempts to investigate questions unaddressed by traditional geography theory. The new economic geographers argue that the agglomeration effect plays a determining role in influencing the spatial distribution of industry and other economic activities. The agglomeration effect leads to industrial concentration in a very limited number of the core regions, and causes regional industrial and economic divergence. Ottaviano and Puga (1997) argue the NEG developed by Krugman and other new economic geographers, has offered a novel approach to answer the sorts of geographical questions which the traditional economic geographers failed to address. They argue that the NEG can illustrate why initially similar or even identical regions can become very different due to the effects of economic agglomeration. Martin and Ottaviano point out that, "Circular causation arises between growth and agglomeration: growth brings agglomeration that fosters growth" (2001, p. 966). Martin and Ottaviano (2001) also clearly suggest that spatial agglomeration and economic growth reinforce each other in a parallel working process.

Martin and Ottaviano's study needs to be taken seriously. Their research suggests that agglomeration effect and regional economic growth go hand in hand with the decreasing costs of transportation and other production. The relationship between agglomeration and economic growth demonstrates causal correlation: they reinforce each other. The agglomeration effect brings huge benefits to regional economic growth; in turn, economic development encourages further regional concentration of industries and economic activities. Due to the increasing returns to economic scale, Rosenthal and Strange's study (2002) suggests that economic agglomeration can bring benefits to entrepreneurs from producing manufacturing goods and engaging in other economic activities at low cost. In terms of this aspect, Brakman and Garretsen (2003) claim the importance of the NEG, whose main characteristic is agglomeration, in understanding regional economic development.

The agglomeration effect is the key to understanding spatial industrial distribution and regional economic development; moreover, business location decisions throughout the world, are also largely affected by economic agglomeration. Hilber and Voicu's logit model test (2005) found a strong relationship between foreign firms' location decisions and economic agglomeration in Romania. Harris's study shows the striking correlation between the geographical location of the American manufacturing belt and regional industrial agglomeration. This manufacturing belt is located in the northeastern part of the USA: which occupies less than 10 percent of the country's total area. However, 70 percent of the manufacturing labour force and 50 percent of

the American market were clustered in this manufacturing belt in the early 1950s (Harris, 1954).

By carefully analyzing numerous case studies worldwide, Porter (1998) highlights the striking correlation between regional agglomeration and spatial economic distribution. Porter points out that the secondary and tertiary industries of Italy are highly concentrated in a few places. For example, the woollen textile industry is clustered in a few regions of Italy. Porter (1998) further states many other cases of regional industrial cluster in the world, including London (British auction industry), Basel (Swiss pharmaceutical industry) and New York (American advertising industry).

4.2.2 History and Agglomeration

The initial factors causing industrial concentration and regional economic agglomeration remain largely unknown. They are probably triggered by both historical and non-historical factors. Ottaviano and Puga (1997) argue that the historical factors can create the self-reinforcing style of agglomeration. In addition, they claim that people's expectation will trigger this agglomeration process. Fujita and Thisse's study (1996) suggests that the prediction of regional agglomeration is difficult. Historical accidents and other initial conditions are important in explaining industrial agglomeration and spatial economic concentration. Fujita and Thisse (1996, p. 371) state that,

This is because the agglomeration of economic activities has the nature of a cumulative, self-reinforcing process and because the emergence of a particular site as a major agglomeration does not only depend upon the intrinsic features of this site. In other words, history matters for economic geography ...

Significantly, Krugman's research into the correlation between historical factors and agglomeration is influential. His is probably one of the most detailed studies on the role of historical factors in causing regional agglomeration and spatial inequality between the core and periphery. Krugman (1991a; 1991b) has researched the relationship between the increasing returns to economic scale and historical factors. Krugman (1991a) believes that the root of increasing economic return and the cumulative effect of economic factors in a few core regions, which directly cause regional economic agglomeration, lies in history. According to the argument made by Krugman (1991a), the increasing scale of economies and associated externalities have significantly influenced the geographic agglomeration of industries; the superior externalities of one region substantially enhance the competitiveness of that place. For example, in Brazil, due to increasing economic return, many factors of economic production have been concentrated in the southern region. At the same time, the numerous intermediate-input producers and large consumer markets are also concentrated here (Krugman, 1999). As far as the USA is concerned, Krugman has

done illuminating work on the formation and development of the American manufacturing belt. Krugman (1991a) stresses that the source of regional agglomeration of the manufacturing industries in the north-east manufacturing belt lies in the increasing economic returns.

Krugman (1991a) firmly argues that the increasing returns to economic scale largely account for the historical factors in analyzing regional economic agglomeration and spatial inequality. The increasing economic returns, regional agglomeration and historical incidents are closely correlated and mutually self-reinforcing. Krugman (1991a, p. 10) stresses that,

In particular, I want to show two things: that increasing returns are in fact a pervasive influence on the economy, and that these increasing returns give a decisive role to history in determining the geography of real economies.

The initial historical incidents are not the direct factor leading to regional economic agglomeration, but the cumulative economic process makes these historical events generate long-term effects on regional industrial agglomeration (Krugman, 1991a). He has cited the cases of prosperous American industries to support his argument regarding the long-term historical impact on regional agglomeration; for example, the regional agglomeration of musical wind instruments in Elkhart, Indiana State and the concentration of high-tech industries on Route 128. Silicon Valley is another example illustrating the prominent historical effect on agglomeration of high-tech industries and venture capital enterprises. According to the argument made by Krugman, due to the original initiative of senior members of Stanford University and the University itself, many high-tech industries, technology enterprises and research centres have clustered in Silicon Valley (Krugman, 1991a).

In his classic book, *Geography and Trade*, Krugman (1991a) argues that the rise of Dalton as the carpet capital of the U.S.A is closely linked to a woman named Catherina Evans. A bedspread handcrafted by Miss Evans was given to their neighbours as a gift. Due to the quality of this bedspread and the obvious skill involved, Miss Evans and her friends became renowned. They launched a handicraft and carpet business, and later dramatically expanded to the other surrounding areas. During the post-World War II period, technological innovation enabled Dalton's carpet industry to enter the weaving machine age and be further developed. Krugman's citing of Dalton shows the importance of historical incidents in affecting regional industrial agglomeration. Moreover, Krugman (1991a) points out another similar agglomeration case of a shoe enterprise in Massachusetts linked with John Dagr, who was a Welsh cobbler. Krugman (1991a) clearly suggests that the emergence of the north-eastern manufacturing belt was closely linked to historical accidents, and seems to firmly believe that the various industrial agglomerations across America should be mainly attributed to such accidents. As he further argues,

The resulting pattern may be determined by underlying resources and technology at some very aggregative level; but at ground level there is a striking role for history and accident. (1991a, p. 66-67)

From my point of view, the influence of historical factors in development of industrial clusters should not be neglected and the significance of history in affecting regional convergence needs to be carefully addressed. However, Krugman's historical argument is not without criticism. His claim regarding the fundamental position of historical accidents in explaining regional industrial agglomeration has certain limitations. Moreover, it is difficult to fully demonstrate the complicated process of regional agglomeration and its interior functioning mechanism. Is it really justified to argue that history is the decisive factor contributing to regional agglomeration? I am not convinced by this argument, as world evidence does not fully support Krugman's historical claim.

For example, the paper and printing technology was originally invented in Ancient China, but the paper and printing industries were further developed and commercialized in the Arabic and Western countries rather than China. Moreover, Wuhan municipality became a crucial trade and transport centre along the Yangtze River region in Ancient China. If history is really the key factor causing regional agglomeration, why has foreign trade not developed and further flourished in Wuhan over the past two decades? Why did the economic centre of the Yangtze River not locate in Wuhan but in Shanghai? The case studies of American industrial agglomeration by Krugman do not represent the whole world and are not sufficient to support his historical view. In fact, even Krugman (1991a) objectively admits that the increasing returns to economic scale and cumulative processes of economic factors should be identified in analyzing regional processes of agglomeration. Rather than, history, these are the key factors which generate long-term effects on the regional agglomeration process. Krugman seems to be overconfident on his historical argument. As he inappropriately states,

... the idea that an economy's form is largely shaped by historical contingency is not a metaphysical hypothesis; it is simply the obvious truth. (1991a, p. 100)

4.2.3 Factor Concentration, Agglomeration and Economic Growth

The clustering of various economic factors reinforces and further develops regional agglomeration. Porter (1998) argues that the effects of scale economies and other external economies tend to make economic factors concentrate in the same region: bringing huge benefits for private firms by providing a better environment for technology diffusion and information spillover. Firms are motivated to locate in that region, which encourages economic and industrial agglomeration. In particular, he highlights the spatial concentration of research and innovation institutions.

Porter's study demonstrates the case of London and its surrounding metropolitan region to support his argument of the important role played by concentration of economic factors in shaping the tendency of industrial convergence. London is the most dynamic and developed place in Britain. The large markets with their huge demand for the various goods and services, and the high concentration of skilled labour, all make London attractive to various industries and firms in terms of economic agglomeration. London is not only the political capital of United Kingdom; but also the British economic and cultural centre. One sixth of Britain's population works and resides in this dynamic municipality and its surrounding satellite regions. Hoggart's study (1991) highlights the significance of London to Britain.

Numerous new economic geographers share Porter's arguments about the effect of economic clustering on regional agglomeration. For example, Krugman's study (1999) suggests that the determinants of spatial convergence and economic agglomeration lie in the balance between the centripetal and centrifugal forces. If the centripetal forces are stronger than the centrifugal forces, regional agglomeration will occur and then be sustained; otherwise, it will cause regional industry dispersion and de-agglomeration. According to the NEG, there are various economic factors (centripetal forces) contributing to agglomeration like transport cost, backward and forward linkages, market, capital, and information. A general discussion of this economic relationship will be offered in the following sections.

4.2.3.1 Transportation Cost and Agglomeration

New economic geographers argue that one of the key factors contributing to agglomeration is lowering transportation costs. The continual fall of transport costs tends to attract firms to locate in the same region: therefore it is the first key factor encouraging industrial agglomeration. Commercial agents can benefit from low transportation costs and good accessibility to large markets, and this has generated huge impacts on spatial industrial distribution. As a result, the regional convergence of industries and economic activities naturally emerges. Vickerman, Spieckermann and Wegenev's study (1995) argues that the partial reduction of transportation cost makes the companies concentrate in the core regions and leads to regional agglomeration. Fujita and Thisse's study (1996) supports the argument made by Vickerman, et al. They point out the tendency of regional agglomeration generated by low transportation cost.

In Portugal, Guimarães, Figueiredo, and Woodward's empirical study (2000) provides evidence for the negative relationship between regional agglomeration of foreign capital, distance and transportation cost. They found that the foreign capital tends to concentrate in the same core region or in regions with a short distance and low transportation cost to the core. Mori and Koji (2001) discuss the positive correlation between rapid economic development and transportation nodes in the Asian economies, such as Hong Kong (China) and Singapore.

4.2.3.2 The Backward and Forward Linkages and Regional Agglomeration

The model of backward and forward linkage is influential to study of geographical industrial agglomeration and regional industrial development. This model needs to be considered seriously. This is because backward and forward linkages not only properly answer the question of spatial industrial concentration: but also more significantly, it indicates the process of regional industrial development. On the one hand, the effect of backward (demand) linkage tends to attract the upstream industries to the region with downstream industry clustering. Located in this place, the upstream industries can better serve the huge market demand of the downstream industries through good accessibility and low transportation cost. Regions with high demand for the intermediate products tend to attract the intermediate-goods producers to agglomerate there. On the other hand, the strong forward (cost) linkage effect makes the downstream industries relocate toward the region which has many upstream industries. In addition to the benefits from the cheap price of intermediate-inputs, due to the competition effect, manufacture of many products of the same type provides more choices for the downstream industries. The effects of backward-forward pattern linkages have strong power in contributing to the unequal industrial distribution and further strengthen regional industrial convergence (e.g. Amiti, 1998 and Henderson, Shalizi and Venables, 2001).

An early study by Hirschman (1959), argued that the linkage effects of two industries' concentration in one place are larger than the effect of these two industries located in separate places. Venables is one of the main contributors to this model of backward and forward linkage. Venables probably provides the best argument regarding the effect of backward and forward linkages in shaping industrial agglomeration. Venables suggests that the cost and demand linkages of the intermediate inputs between the upstream and downstream firms lead to regional agglomeration. He points out that the effect of these linkages can largely affect the firms' location decisions. Venables (1996, p. 341-42) states that, "Putting the demand linkage and the cost linkage together creates a force for the agglomeration of activity in a single location."

Due to transportation cost reduction, the studies by both Krugman and Venables (1996) and Walz (1996) suggest that the cost and demand linkages between the intermediate and final product firms can cause regional concentration, and form the core-periphery pattern of spatial disparity. As far as this pattern is concerned, the different regions will not demonstrate a uniform pace of economic growth. Puga and Venables's study (1996) provide the analysis of cause of cost and demand linkages. They further highlight that the backward and forward linkages of the intermediate inputs' supply and demand would make the industrial enterprises concentrate in a single region in order to enjoy the benefit of pecuniary externalities. Amiti's study (1998) suggests the economic benefits brought by this demand and cost linkages for both the companies engaging in the upstream or downstream industries. In particular, for the companies, Amiti points out that the main advantages of this linkage are

generated from transportation cost reduction, large market impact of the downstream industry on the upstream companies, and the benefit of the upstream firms' competition for the downstream enterprises.

One of the main benefits contributed by the backward and forward linkage is probably generated from the competition effect in the upstream industry. Competition means improvement and competition indicates development. The concentration of the rival upstream industries in a region can significantly stimulate the firms' competition motivation, force them to improve and enhance their production efficiency. In addition, the rival competition brings knowledge spillover and other external economies which benefit all of the companies. Porter's research (1998) highlights the benefits of competition in attracting the downstream and upstream firms to locate close to each other. He argues that the competition effect can boost the firms' productivity and help them to take advantage of external economic activities, such as innovation, training and knowledge spillover. All of the upstream and downstream industries might concentrate in a few regions due to this kind of linkage. Therefore, consequently, it leads to the tendency of regional economic concentration. Henderson, Shalizi and Venables (2001) stress that the regional agglomeration of both upstream and downstream industries is mainly caused by the linkage effects. As they maintain,

There is thus a positive feedback between location decisions of upstream and downstream firms, tending to draw both types of firms together in the same location, so leading to agglomeration. (2001, p. 83)

There is widespread evidence illustrating the importance of backward and forward linkages to regional agglomeration. In France, Combes and Lafourcade's study (2002) suggests that the effect of backward-forward linkage is one of the key factors in spatial convergence and regional concentration of economic and industrial activities. Combes and Lafourcade (2002) maintain that concentration of the upstream intermediate-input producers forces these companies to enhance productivity and lower product price. As a result of these effects, the downstream consumers have strong incentives to relocate to the same or surrounding areas in which many upstream companies are located. Due to the effect of backward and forward linkages, Shalizi and Venables (2001) found that the heavy and transport industries in South Korea were clustered in a limited number of large cities. Furthermore, in Britain, Cambridge is a good example to demonstrate the linkage effects of backward and forward industries in influencing regional convergence of technology innovation and research activities. Equally, this argument also holds true of the agglomeration case, Silicon Valley in the U.S.A.

4.2.3.3 The Market Factor and Agglomeration

In a world characterized both by increasing returns and by transportation costs, there will obviously be an incentive to concentrate production of a good near its largest market, even if there is some demand for the good

elsewhere. (Krugman, 1980, p. 955)

The NEG indicates that the market factor can dramatically stimulate the spatial concentration of firms and manufacturing industries. Therefore, this large market effect causes regional industrial clustering. Krugman is one of the most respected scholars in developing the NEG. His model of the market effect is influential in analyzing regional economic agglomeration. Krugman's model suggests that the individual commercial agents have strong incentives to cluster in the regions near to the large markets. This is because these commercial companies can benefit substantially from the increasing returns to economic scale and low transport cost generated from the large market effect. Locating in the core region with a large market, he argues that both the firms engaging in the intermediate- and final-goods production benefit from increasing economic returns and huge demand. The market effects tend to attract numerous industries to concentrate in the same region, including manufacturing and service industries. Krugman's model is widely supported by other new economic geographers (e.g. Davis and Weinstein, 1999), Fujita and Thisse (1996) and Ottaviano and Puga (1997), in particular, they highlight the determining role of large regions with big population size in generating scale economies and reducing transportation costs. The influence of large markets in the industrial structure and economic development of the urban region is decisive. For example, Ottaviano and Puga's study (1997) offers a clear statement regarding the role played by the large market in contributing to regional economic agglomeration. They state that,

However, the combination of increasing returns to scale and trade costs encourages firms to locate close to large markets, which in turn are those with relatively many firms. This creates pecuniary externalities which favor the agglomeration of economic activities. (1997, p. 724)

Henderson, Shalizi and Venables's research (2001) further suggests the incentives for the private firms clustering in the same region contributed by the large market effect. In terms of market-centralized agglomeration, an important benefit is the scale economies, which provide the strong incentives for companies to concentrate in the regions near to the big markets. In terms of this aspect, the new economic geographers, Fujita and Thisse (1996) stress that the increasing returns to economic scale are the key factor to understanding spatial industrial distribution and regional convergence of the other economic activities.

However, first, the question of the emergence of large market should be addressed. How has this market been developed? What is the formation process of a large consumer market? My understanding is as following. On the one hand, the geographical factors are crucial in the development of large market. In precise terms, inherited location advantage has provided a good environment for big market development. In particular, the coastal regions or the areas near to the river tend to be favoured in development of markets. This is possibly because riverside land was generally fertile and easy to irrigate, which then enabled development of a prosperous

agricultural sector. Therefore, people tended to move to these regions: where they could easily make a living and avoid food shortages. The Pearl River and Lower Yangtze River Deltas are two classic examples, where the emergence of large consumer market was closely associated with fertile land and rich water resources. On the other hand, history has also played an important role in market formation. Regional development history indicates that almost all of the regions with large markets have inherited the glorious development history of civilization, trade and culture; for example, London, Beijing, Shanghai and Paris. As a result of these two factors, the total population size enlarged and large markets have steadily developed in these regions.

In terms of the European Union (EU), accompanying its rapid economic integration, many manufacturing industries have concentrated in a limited number of regions, due to the effect of scale economies, and some European countries have become specialized in one particular kind of product manufacture (Amiti, 1998). Glaeser, et al, (2001) use the case of restaurant concentration to support the contribution of economic scale to regional agglomeration. Rosés's case study of Spain (2003) suggests that unequal industrial distribution between the northeast, central and northwest areas is largely shaped by the large home market effect. Moreover, Rosés argues that the market effect and agglomerative economies are the determining factors in explaining the spatial income disparity and regional divergence within Spain. World evidence demonstrates that the market effect is crucial in influencing regional distribution of industrial and other economic activities, and unbalanced regional economic development within a country. For example, Friedmann's study of Venezuela (1966) is a good case to illustrate the market's effect on spatial convergence of economic activities and regional industrial agglomeration. Many of Venezuela's manufacturing industries are concentrated in the Caracas – Valencia region. The high share percentage of national population and national wealth concentration within the Caracas – Valencia region, and, the large percentage of national population and wealth concentrated within this region provide a large market for Venezuela's manufacturing companies.

In Japan, Davis and Weinstein (1998) have offered case analyses to argue the reciprocal relationship between market and agglomeration by demonstrating that manufacturing industries have agglomerated in certain cities and prefectures within Japan due to the market effect. The attractiveness of large consumer markets tends to encourage private firms to concentrate in the core metropolitan regions. Close to these core regions, the enterprises can take advantage of the huge consumer demand and serve the majority of consumers with low transport costs. Amiti (1998) raises the cases of central European countries like Poland, Romania and Czech Republic. Because they have good transport access and are close to large western European markets, these central European countries have attracted many industries. Moreover, Hanson's study of Canada has pointed out the role of foreign market in contributing to regional agglomeration. Hanson (1998) argues that the proximity to both the large American and Canadian car markets have made the Michigan – Ontario industrial

corridor the car industry centre. Hamilton's study (1991) argues that London, with its own large market and the inherited advantages in accessing other foreign markets (e.g. EU), has attracted many kinds of firms to locate their headquarters and offices there.

In the context of the U.S.A, the financial services industry is highly concentrated in New York mainly due to the market effect. In terms of the financial companies, the attractiveness of doing business in New York is that many clients and skilled human talents specializing in financial business concentrate in New York (Krugman, 1999). In addition, Hanson's study (2005) provides evidence of striking correlation between firms' location decision and market effect. In order to enjoy the benefits of increasing returns to economic scale and low transportation cost, the private agents prefer relocating to regions with large consumer markets. Hanson found that the correlation between spatial industrial agglomeration and market linkage effect was highly positive and prominent. Harris's study (1954) illustrates that many industries are concentrated in the regions with large consumer markets e.g. the American automobile and agricultural industries.

In Europe, the emergence of regional agglomeration tends to focus on the rich and developed regions with large markets. Krugman and Venables (1990) argue that the most prosperous and developed regions are concentrated in the north of Western Europe. Davis and Weinstein's study (1998; 1999) also suggests the case studies of the OECD and Japan to demonstrate the importance of the market effect. Hanson's study (1998) illustrates that many of Mexico's industrial centres have relocated to the northern cities, such as Ciudad Juarez, Monterrey and Tijuana from Mexico City. This movement is because of geographical proximity to the American consumer markets. As a result, Hanson found that Mexico City's share in the total manufacturing employment of Mexico had dramatically decreased during the period from 1980 to 1993.

4.2.3.4 Effects of Capital and Labour Concentration on Regional Agglomeration

Myrdal (1957) points out the uneven capital distribution between the core and peripheral regions. Due to the greater commercial opportunities and higher investment returns, Myrdal indicates that capital tends to move toward the developed core region rather than the under-developed periphery. In addition, many other scholars (e.g. Williamson, 1975; Wang and Hu, 1999) argue that capital concentrates on the more developed and prosperous regions for reasons such as low investment risk, entrepreneurial ability and the efficiently functioning capital market. As a result, capital concentration leads to widening spatial disparity. Friedmann's research (1966) reaches similar conclusion about the unequal capital distribution. Friedmann (1966, p. 22) argued that:

Entrepreneurship and capital will tend to be heavily concentrated in the more highly developed regions and will be attracted into new areas to the extent that profitable opportunities are perceived.

The unequal capital distribution between the rich and poor regions illustrates that, due to the effect of high investment returns, capital tends to agglomerate in the developed regions, and this is one of the important factors shaping economic agglomeration. In terms of the poor periphery, a shortage of capital could cause it to remain in the dilemma of growth stagnation and economic backwardness. Lack of investment capital might further marginalize the small economy of this region: in the end, inevitably contributing to regional income inequality.

As far as labour is concerned, because of the generally higher wages and greater work opportunities in the developed regions, many native workers from the poor periphery migrate into the rich metropolitan regions in order to make a better living. Consequently, this unequal distribution of workers, causes labour shortage in the peripheral region, as described by Wang and Hu (1999).

For companies located in the core regions, labour agglomeration can bring enormous benefits in terms of reducing labour costs and staff recruitment time. In addition, labour agglomeration directly contributes to rapid economic and industrial development in the rich and developed regions. Kaldor (1996) argues the positive influence of human talent concentration in stimulating industrial clustering. For example, undoubtedly, industrial and economic agglomeration in London has benefited substantially from high concentration of labour: industrial workers, talented individuals or people with specialized skills, all concentrate in the Greater London area. Hamilton's study (1991) provide a detailed analysis regarding labor concentration contribution to the economic and industrial development in London.

However, to be fair, it is important to recognize that the strong effect of regional labour agglomeration on widening the spatial economic disparity between the core and peripheral regions is not the whole story. Migrant workers from the poor peripheral regions have significantly contributed to rapid economic growth in the developed regions; and this trend needs to be analyzed. China is a very good example of the migrant worker issue due to the large numbers of the return of migrant workers to their under-developed hometowns in China. Three representative hill counties, located in Jiangxi province, were chosen for the field research by Murphy (2002). According to her estimation, about one third of migrant workers who come from economically poor regions in China have returned to their hometown since 1995. Reasons for their return includes: family obligation, city discrimination, personal illness and urban unemployment. A pull-push pattern of incentive factors has been developed by Murphy to assess the causes of migrant workers returning to their backward hometown²².

²² "A push – pull perspective can also be used to explain return migration. As examples, 'push' factors would include job insecurity, poor living conditions, social discrimination, and legal restrictions on urban residence. 'Pull' factors refer to expanded employment or investment

Murphy has provided information about the various contributions made by the returning migration workers to their under-developed rural hometowns. The local authorities in Jiangxi seem to firmly believe in the benefits brought by the migration workers to economic development in this province. In particular, their financial contribution has been highlighted by Murphy's study. She argues the importance of remittances made by the returning workers in supporting their poor families and hometown.

According to Murphy's fieldwork (2002), in Jiangxi province, remittances have played an important role in the rise of rural income per capita. In 1992, nearly 50 percent of the rural income increase in Jiangxi could be attributed to the remittances sent by the returning migration workers. In addition, the new market information and advanced skills brought back by the returning workers have contributed to economic and social development in the poor interior regions within China. Murphy (2002, p. 111-12) suggests that,

Remittances provide funds needed for maintaining the social customs associated with house building and marriage, while migration changes the values and relationships that social actors reproduce through these customs. These changes result not only from modernization or diffusion but also from more fundamental transformations occurring within rural society.

From my point of view, regional labour concentration is one of the indispensable factors for economic agglomeration. In the under-developed periphery, the slow industrial and economic development caused by a shortage of human capital has negatively affected regional competitiveness and limited potential for future economic growth. The contribution (financial and non-financial) to the economic growth in the peripheral regions made by the returning migrant workers should not be overestimated. Murphy's field study of three selected counties located in Jiangxi is spatially biased; and the interview sample is obviously too small and unrepresentative. She seems to be a little over-optimistic regarding the contribution made by the returning migrant workers to their native hometown. First, workers returning due to injury or ill-health might put huge financial pressure on their families and the health care system of their hometown. Secondly, the increase in violent crime perpetrated by unemployed returning workers (although the exact crime number remains unknown), might have brought much disruption to the political stability and social harmony of the poor regions, which might further discourage incentives to potential foreign investors. Thirdly, readjustment problems in the rural hometown for the returning migrant workers might also have a negative effect on the economic growth in the peripheral region. Last but not least, the significant number of unemployed people in

opportunities in the origin areas, access to land or opportunities for acquiring property, high labor demands sustained by low-technology cultivation methods, and the presence of family and kin." (Murphy, 2002, p. 19)

the under-developed periphery might be a side-product of migrant workers' return: as there may not be enough job opportunities for this increasing amount of returning population. Significantly, these unemployed returning workers have made little positive contribution to development in the poor regions. Relative to the developed coastal area, the massive rural and interior regions of western China seem to be largely untouched by the rapid industrialization and dramatic national economic development of the reform period. Due to a lack of working opportunities in the periphery, many returning workers have subsequently re-migrated into the developed core regions. In Jiangxi, Murphy's field research (2002) has pointed out the hardship among the returning labourers, their inability to readjust to rural life and frustration with the society of the peripheral region. Murphy (2002, p. 217-18) has made points about the negative effects of the returning workers on the under-developed regions,

First, not all migrants make money in the cities: unemployment, poor working conditions, injury, and illness commonly afflict migrant workers. Second, some households are adversely affected by the ways in which resources and burdens are distributed within the family.

Therefore, the significance of returning migration workers in boosting local economic growth in the economic poor regions should not be overestimated. The main development theme is perhaps the widening of spatial economic inequality between the core and periphery contributed by the migrant workers.

4.2.3.5 Technology Spillover and Regional Agglomeration

The NEG demonstrates that regional agglomeration brings benefits for the manufacturing and service firms through the spillover of knowledge and concentration of technological innovation. Moreover, it suggests that regional agglomeration and diffusion of technology reinforce each other. The agglomeration effect tends to encourage technological information and innovation activities concentrate in a few core regions; in turn, information concentration and innovation spillover further strengthen regional agglomeration. Martin and Ottaviano (2001) claim the causal relationship between technological innovation and agglomeration. As they suggest,

Economic agglomeration in one region spurs growth because it reduces the cost of innovation in that region through a pecuniary externality due to transaction costs. (2001, p.947)

The costs of technological diffusion and knowledge spillover among individual firms are lower when all of the firms locate in the same region. Porter (1998) stresses that geographical industrial convergence speeds up the rate of information flow and innovation diffusion. Keller's study (2002) suggests that the farther away a country is from the information and knowledge centre, the less it benefits from technology spillover. For example, in contrast to Australia, Netherlands can enjoy up to 40

percent more in terms of positive effects of R&D spillover generated from the G-5 countries (France, the United Kingdom, the United States, Germany, and Japan). Keller (2002, p. 120) claims that, “I find that technology is to a substantial degree local, not global, as the benefits from spillover are declining with distance.”

Rosenthal and Strange (2001) further argue that the innovation- and knowledge-oriented firms concentrate in the same region in order to share information and intermediate input. For example, in Wall Street, New York and the Silicon Valley, Glaeser, Kolko and Saiz (2001) demonstrate the benefits derived from spatial proximity to advanced information and mutual information sharing.

Concentration of economic factors leads to the tendency of economic agglomeration in the core regions. The economic cluster factor is not only the key factor affecting the regional agglomeration of the industrial and other economic activities; but also more significantly, it plays a determining role in affecting spatial industrial distribution and regional economic growth. Friedmann (1966) points out the benefits for firms generated from the concentration of economic factors. Such concentrations within a region tend to attract the private firms and industries locate there. Consequently, it causes the regional convergence of industries and economic activities. The convergence of all of these economic forces in the core regions strengthens the effect of regional agglomeration; by contrast, the agglomeration effects will further encourage regional concentration of economic forces. As a result, the relationship with the reciprocal characteristic between agglomeration and economic forces concentration attributes to the emergence of the self-reinforcing pattern of regional convergence. This topic is discussed in detail in the following section.

4.2.4 Opposing Forces to Agglomeration

Although there are numerous factors contributing to regional economic agglomeration, some factors act as opposing forces (centrifugal forces) against agglomeration. Under these circumstances, they might lead to regional de-agglomeration. For example, strong opposing forces, like traffic congestion can negatively affect the tendency of regional agglomeration. Therefore, in order to provide an objective and accurate estimation of regional agglomeration, both the benefits and costs of economic agglomeration need to be addressed. Parr (2002) points out that the costs brought by regional economic concentration should be estimated in order to evaluate the net value of economic agglomeration. Martin and Sunley’s research (2003) details the potential advantages and disadvantages to private agents which are caused by regional industrial convergence (see Table 4.1 below).

Table 4.1: Benefits and Costs of Industrial Clusters

Claimed advantages	Potential disadvantages
Higher innovation	Technological isomorphism
Higher growth	Labor cost inflation

Higher productivity	Inflation of land and housing costs
Increased profitability	Widening of income disparities
Increased competitiveness	Over-specialization
Higher new firm formation	Institutional and industrial lock-in
High job growth	Local congestion and environmental pressure

Reproduced from Martin and Sunley (2003, pp. 27)

An early study, Hermansen (cited in Kuklinski, 1972) stresses that there are important opposition factors to regional agglomeration, such as traffic congestion, crime and pollution. Henderson, Shalizi and Venables's research (2001) also suggests that market de-centralization and dispersed consumer demand can cause regional industrial divergence.

4.2.4.1 Traffic Congestion

The negative forces opposing regional convergence can have a large influence in reshaping the economic and industrial agglomeration proposed by the new economic geographers. Under certain extreme circumstances, these opposite factors might force the firms and industries to relocate to the other peripheral regions, and lead to a slow tendency of industrial dispersion and regional de-agglomeration. This is especially true of traffic congestion. Owen (1987) points out the cases of traffic congestion in Thailand and their associated negative effects. Ten percent of Thais (around 5 billion) are residing in Bangkok and its surrounding metropolitan region; while, 25 percent of Thailand's GDP has been produced in this area. He indicates that regional economic convergence has caused severe problems of traffic congestion. Moreover, Owen (1987) raises the Brazilian case. During the period between 1940 and 1989, Brazil had made impressive progress in terms of urbanization: by the end of the 1980s, about two-thirds of Brazilian people lived in the cities. In contrast to the equivalent figure in 1940, the proportion of urban population had multiplied twice. An urban population of 75 million has agglomerated in a few core regions, such as Rio de Janeiro and São Paulo, and this has undermined the original advantages of regional economic convergence. In general, Owen (1987, p. 74) states the serious impacts of traffic congestion brought by economic convergence,

In Latin America, a heavy concentration of population and economic activity in a few very large metropolitan cities has created massive traffic jams that detract from living conditions and increase production costs in the industrial sector.

Many other regions in the world have been facing similar traffic congestion challenges. Traffic congestion costs have had an extremely negative impact on regional economic growth and spatial industrial concentration. Owing to economic agglomeration, the developed and core regions have already gained large amounts of various industries, firms, population and market. The impact of economic agglomeration tends to further attract more industries, companies, migration workers

and investors to move toward these core regions, which inevitably will bring the problem of heavy traffic congestion. Relative to the peripheral regions, the companies locating near to the core regions have to pay much higher transport cost and suffer from long delays on the road. The agglomeration benefits to the regions might be partly or even entirely offset by heavy congestion costs. Kraft, Meyer and Valette's study (1971) illustrates the American case to demonstrate the negative effect of congestion.

Because of the serious problems of these opposite forces, many firms did not want to locate in the core and developed regions. This indicates why not all of the firms and industries concentrate in one region, even if the economic agglomeration effect is very powerful. For example, American high-tech industries are concentrated in Silicon Valley rather than New York, Chicago or San Francisco. Even though these cities have large consumer markets, they suffer from high property prices, high average wages and traffic congestion. In terms of China, Beijing and Shanghai are the two largest municipalities; however, it is certainly not the case that all of the Chinese firms and industries have clustered in these two places. For example, in China, Dongguan and Kunshan are two important manufacturing centres of computer and other IT products. Shaoxing is one of the key production bases for the textile industry.

4.2.4.2 Transportation Cost Reduction

Low transportation cost is another important factor opposite to regional agglomeration. Glaeser and Kohlhase's research (2004) strongly suggests that the industrial concentration in the core region and regional agglomeration will be shaken and disappear eventually, due to the shrinking importance of infrastructure centres and the impact of dramatic reduction of transportation cost. Due to the impact of transportation cost elimination, Glaeser and Kohlhase (2004) highly doubt the self-organizing pattern of regional agglomeration and the persistent regional convergence claimed by the new economic geographers. They argue that the credibility of the NEG needs to be reconsidered and updating of Krugman – Fujita – Thisse's agglomeration model is badly needed. Glaeser and Kohlhase seem to claim that the NEG model is only suitable for past analysis of regional economic convergence. An updated model with the absence of transportation cost can be better analyzed regional development in future.

However, in terms of transportation cost reduction, in my view, Glaeser and Kohlhase' argument is not so convincing. The previous chapter has demonstrated that transportation cost would not completely disappear, even if the power of new technological innovation impacts on manufacturing production, information cost and the way of business communication. As discussed earlier, many kinds of business, service and communication cannot be completely finished without the face-to-face interaction and tangible production transportation. Therefore, the involvement of distance and transportation costs in regional economy analysis cannot be denied. Many previous researches have argued this point. For example, Combes and

Lafourcade's study (2005), based in France, has criticized the argument made by Glaeser and Kohlhase, and stressed that transportation costs will persist for a long time.

To be fair, some opposite forces have brought challenges to the spatial industrial concentration and economic convergence in the core regions. However, it is important to recognize that the majority of these core regions still remain dynamic and economically prosperous, and continue to maintain fast economic development. More significantly, the tendency of regional economic convergence has not been interrupted to any appreciable extent. As far as the self-reinforcing nature of agglomeration effect is concerned, it clearly demonstrates that the centripetal forces which favour regional agglomeration are much stronger than the centrifugal forces. If the economic benefits generated from the agglomeration effect are higher than its associated costs, the private enterprises are willing to tolerate and accept these opposite forces to agglomeration. For example, in terms of the firms located in the core regions, they are willing to pay high transportation costs due to the substantial benefits brought by regional agglomeration. Therefore, the key point is how seriously these opposite forces actually are: whether or not they are powerful enough to change the direction of regional agglomeration. In my opinion, the answer to these questions is 'No'; it seems to be always the case that economic agglomeration is stronger than the effects of its opposing forces. Private agents continue to concentrate in the core regions; and reinforce the leading economic positions of these regions, above all, the tendency of regional agglomeration. The powerful self-reinforcing pattern of agglomeration effect, which continually attracts more and more firms to cluster in the core regions, can largely offset the potential cost of these opposite forces. For example, London has been facing increasing environmental challenges brought about by traffic congestion (Pharoah, 1991). However, London still remains one of the most dynamic regions in Britain and the world. As far as the private companies are concerned, it seems that the economic benefits contributed by proximity to London are much higher than the costs of opposing forces, such as traffic congestion.

4.2.5 The Self-Reinforcing Pattern of Agglomeration

The self-reinforcing (or self-organizing) feature of production tends to generate the 'lock-in' effect on the existing geographic advantages and factor concentration in the core region; it strengthens the competitiveness of the core region and its leading position in economic development. The self-reinforcing process further reinforces the clustering tendency of the economic factors; it also acts as a catalyst in contributing to the sustainability of economic agglomeration in the core region. Numerous new economic geographers claim that the strong self-reinforcing feature of agglomeration effect is the key to sustainable economic development in the core region. For example, Krugman argues that once a location is chosen, the firms tend to stay in that place. His study stresses that industrial agglomeration has a strong sustainability characteristic

once it is established. Krugman (1991, p. 80-81) states that, "Thus there is a circularity that tends to keep a manufacturing core in existence once it is established."

Krugman (1991) demonstrates the case study of the American manufacturing belt to support the self-reinforcing agglomeration model. He argues that the 'lock-in' characteristic of the agglomeration effect reinforces the initial geographical advantage of the manufacturing belt. He seems to believe that the continual prosperity of this manufacturing belt is largely due to the strong effect of self-reinforcing pattern agglomeration. Krugman's model provides a clear argument regarding the self-reinforcing feature of regional production. His study demonstrates that the contribution of economic concentration to agglomeration is further reinforced by an inherited self-reinforcing process. Furthermore, Krugman (1999) indicates that the effect of the self-reinforcing nature of agglomeration is more important than the inherited advantage of geographical location. The self-reinforcing process of agglomeration strengthens the economic development in the developed region, even if their geographical advantages disappear. The research conducted by Gallup, Sachs and Mellinger (1999) highlights that the main difference between the traditional and new economic geography is the effect of self-reinforcing agglomeration in influencing regional economic development. The agglomeration effect can generate the self-reinforcing process of production; it can further strengthen the previous location advantages of the core metropolitan areas. Gallup, Sachs and Mellinger (1999, p. 184) maintain that,

New economic geography models illustrate the possibility of self-organizing spatial patterns of production based on agglomeration effects rather than on differences in climate, transport costs or ecology.

In addition, Henderson, Shalizi and Venables (2001) analyze this self-organizing pattern of spatial agglomeration effect, and provide some useful explanation for the cause of this pattern production; in precise terms, they are the 'lock-in' effect and historical accidents.

4.2.5.1 Path Dependence, History and Agglomeration

The path dependence and historical factors are closely correlated. They have strong impact on regional agglomeration development. The effect of path dependency further strengthens industrial agglomeration once it has occurred. Paul David is a representative scholar of path dependence²³. David (1994) suggests that newly arriving components and other determining factors are very sensitive to the initial situations. The new components and other factors will deliberately adapt themselves to fit the pre-existing economic structure. The new economic factors will tend to

²³ According to David (2000, p. 1), "The concept of path dependence refers to a property of contingent, non-reversible dynamical processes, including a wide array of biological and social processes that can properly be described as 'evolutionary'."

follow existing economic and institutional patterns, and the new components will not try to separate themselves from the existing factors. This is because dramatic or revolutionary changes are too costly. Moreover, the new economic components do not want to create chaos or dysfunctional situations for themselves.

In his later study, except for strong exogenous forces or a large external shock, David (2000) argues that the impact of path dependence will attract all of the private agents to lock in the pre-existing places rather than other newly created regions. Regional economic agglomeration will naturally occur and be further reinforced in the original place. David (2000, p. 10) states that,

Path dependent systems - which have a multiplicity of possible equilibrium among which event - contingent selections can occur - may thus become locked in to attractors that are optimal, or that are just as good as any others in the feasible set, or that take paths leading to places everyone would wish to have been able to avoid, once they have arrived there.

David highlights the historical factors and institutional framework as causing path dependence. The historical factors and incidents are important for the institutions' development. As David claims, "institutions are the carriers of history" (1994, p. 208). The initial situations and historical incidents, which formed and developed the institutions, cannot easily be broken or changed. Due to the effect of path dependence, the private agents, governmental institutions and other industrial organizations tend to agglomerate in the same place. This example of regional industrial cluster can generate the self-strengthening characteristic. According to the argument made by David (1994), this kind of self-reinforcing mechanism, which is closely connected to governmental concentration, tends to further develop regional agglomeration, and make the pre-existing regional cluster of the forward and backward industries become even stronger.

The path dependence research done by David needs to be taken seriously as it has broadened our knowledge of the self-reinforcing pattern of regional agglomeration. David's path dependence theory has made the NEG model of the self-organizing type of agglomeration process more reasonable and credible. In terms of this point, historical factors make a difference to regional economic development. The historical incidents and the pre-existing institutional structure are the two indispensable factors in understanding the correlation between path dependence and regional convergence. In addition, David's study identifies the impact of people's consistent expectation and self-controlled motivation in addressing the path dependence theory and forming regional economic cluster. This finding is also justifiable.

Although he does not directly analyze the issue of path dependence in affecting regional disparity and industrial agglomeration, Krugman's research on regional convergence and NEG seems to be benefited by David's path dependence theory. In

his early study, in the emergence of the industrial cluster and regional economic centre, Krugman (1991a) has already indicated the role of strong path dependence based on the scale economies and transportation cost reduction. The initially favoured conditions of one region can reinforce its industrial agglomeration, and the path dependence effect seems to inhibit any opportunities for the periphery to overtake the existing core region. Therefore, path dependence theory and Krugman's self-reinforcing agglomeration model clearly indicate that regional economic divergence is the inevitable outcome of spatial development.

From my point of view, the argument of the self-organizing pattern agglomeration effect is one of the key contributions made by NEG to the study of modern economic geography. Even if the inherited advantage of favourable geography in one region becomes less important or completely disappears, the NEG model clearly suggests that the economic and industrial development in that region can be sustained, and its economic prosperity can persist in the long term. This is because the power of the self-organizing characteristic and self-reinforcing process of agglomeration is strong enough for core regions to further develop. The evidence demonstrated in many previous studies illustrates the effect of this self-reinforcing process in stimulating economic development. For example, in Mexico, the past economic success of Mexico City is mainly a result of the favoured natural geography. Thanks to the self-reinforcing pattern of economic agglomeration, this city remains prosperous even if the superior locational advantage becomes insignificant (Krugman, 1999). Moreover, Krugman cites the case of São Paulo, Brazil's largest economic centre to further analyze this self-reinforcing spatial pattern of production. In terms of America, Fujita and Mori believe that the port cities remain prosperous due to self-reinforcement even when the sea-based trading activities become less important for the economy of these coastal regions. Their study suggests the key role of self-reinforcing patterns of agglomeration in boosting regional economic development. Fujita and Mori (1996, p. 94) argue that,

If [**American largest cities**] they were 'neoclassical port cities', then they should have disappeared a long time ago when the original advantage (of cheap water access) became unimportant. Clearly, their continued prosperity can be explained only when we consider the 'lock-in effect' of some self-reinforcing agglomeration forces. (Bold content added)

Even when the inherited geographic advantages have become less important, these developed regions within America are still prosperous and achieve further economic development. Moreover, their leading positions of economic growth seem to be highly unshakeable. The American industrial firms still prefer to locate in these dynamic places rather than relocating into the peripheral regions. Therefore, there are sufficient reasons to claim that the self-reinforcing pattern of agglomeration effect might strengthen the concentration tendency of industrial and other economic activities in these coastal places; it further reinforces the existing leading position of

these American cities, even if their favoured geographic location becomes less prominent.

4.2.6 New Economic Geography: its Critics

The NEG has been widely discussed by the academic community. Some scholars have, indeed, raised strong doubts and criticisms of the whole NEG model. Neary's study (2001) suggests that new economic scholars like Krugman and Fujita have not provided any empirical case studies or proposed any useful policies which can be utilized by governments. Similarly, in maintaining that the NEG has suffered from numerous problems, Clark (1998, p. 75) claims that,

Krugman uses stylized facts like the existence of increasing returns to scale, joined together with his analytical methods in a self-reinforcing circle: his chosen stylized facts allow for the application of analytical methods, and those methods, developed in the study of international economics, explain his chosen stylized facts of economic geography.

The critics claim that the NEG is far from perfect, and in fact, has serious shortcomings. Martin and Sunley (2003) have made some strong criticisms of the regional agglomeration benefits and the whole model of the NEG. According to Martin and Sunley (2003), the new economic geographers did not clearly identify the precise linkage between the industrial agglomeration of a region and its economic growth. What is the correlation between regional industrial cluster and economic development? Martin and Sunley suggest that the NEG did not provide a universal or convincing analytical framework to answer this question. Due to a lack of relevant or strong evidence to support the NEG theory, they seem to suspect the major contribution of industrial concentration to regional economic growth.

However, Martin and Sunley's argument is not accurate. To be fair, Krugman and other new economic geographers have provided numerous case studies regarding economic concentration and spatial industrial cluster. For example, Krugman has made perceptive case studies of America and Europe. Besides the rich evidence discussed earlier, Fujita and Thisse (2002) cite the case of France. Paris and its surrounding metropolitan region account for less than three percent of land area of France, but about 30 percent of French GDP has been produced in this small core region. In addition, Fujita and Thisse (2002) offer a very detailed study of regional economic concentration in East Asia; particularly in Japan. Their study suggests that the national wealth and economy of Japan have been concentrated in the core areas within the five prefectures of Tokyo, Kanagawa, Hyogo, Aichi and Osaka. Up to 31 percent of manufacturing employment and 40 percent of total GDP in Japan was generated from these small areas, where the land merely accounts for about five percent of total national land area.

One of the most critical arguments against the NEG is probably from Ron Martin. Martin (1999) argues that this model just reconfirms or re-states the traditional geography and international trade theories. He firmly believes that the NEG is the re-examination of the traditional theory of geographical location. Further, Martin claims that the so-called 'new economic geography' is actually nothing new and certainly can not be regarded as geography. In my view, Martin's study is partially right. It is certainly the case that the NEG model is conceptually based on previous theories of regional economy and geography, such as international trade and traditional location theory. Moreover, the new economic geographers have benefited significantly from the ideas of many previous scholars. In fact, many previous researchers have already partly raised or indicated the important role of regional factor concentration and scale economies in influencing agglomeration. For example, in 1920, Marshall pointed out the importance of industrial concentration and agglomeration in certain manufacturing towns,

The advantages of variety of employment are combined with those of localized industries in some of our manufacturing towns, and this is a chief cause of their continued economic growth. (1920, p. 271, cited in Rosenthal and Strange, 2002, p. 11)

Moreover, Hirschman (1959) argues that the private firms and manufacturing industry consistently overestimated the importance of external economies to the core regions; he referred to it as the 'growth pole'. Due to the attractiveness of external scale economies, Hirschman believes that the private agents and various industries will concentrate in one core region in order to take advantage of the external forces. Hirschman's regional theory will be further considered in detail in the following section. Moreover, the new economic geographers have taken some ideas from Lösch. For instance, in his early study, Lösch (1954) recognized the contribution of agglomeration in affecting the spatial distribution of economic activities. Lösch's study suggested that the presence of a natural geographic advantage (such as port) is not enough for a region to achieve and maintain economic development, or for firms to obtain maximum economic profits. Economic agglomeration tends to attract companies to concentrate in a particular region with a large labour supply, many intermediate-goods producers and large consumer market. The proximity to this region would enable the firms to obtain the maximum profits via production cost reduction and the spillover of scale economies.

Discussion of previous geographic studies, in fact, illustrates that the new economic geographers have benefited from the previous researches by Marshall, Hirschman and Lösch. These scholars have inspired many good ideas for the new economic geographers. To an extent, some arguments claimed by the new economic geographers are established on the previous research. The NEG model is developed under the foundation of previous geography theory and other disciplinary studies. In fact, Fujita and Thisse (1996) admit this point.

However, in my view, it is unfair and inappropriate to claim that the NEG just re-investigates or re-works previous studies, and offers no new knowledge to modern economic geography. The NEG has systematized the previous research; presenting detailed frameworks for mathematical analysis. More significantly, the new economic geographers have provided a new and developed geographic model to analyze spatial economic distribution and regional industrial agglomeration. The key research findings suggested by the new economic geographers are perhaps the increasing returns to external economies scale and self-reinforcing pattern of regional agglomeration. This model helps scholars to further understand the complicated process of regional economic development. For example, the NEG not only has proposed the model of the forward and backward linkages; but also it suggests the self-reinforcing pattern of agglomeration effect. Krugman and Venables (1995, p. 859) argue that,

... our model exhibits behavior different from that of either antecedent: the interaction between transport costs and trade in intermediates creates country-specific external economies, which may lead to agglomeration of industrial activity.

In terms of this aspect, Brakman and Garretsen (2003) offer relatively fair judgment of the NEG. They claim that the new economic geographers provide a meaningful starting analysis of regional economy, and this is the only one mainstream economic theory which systematically addresses the issue of geographic location and economic concentration.

Another critical argument raised by Martin (1999) suggests that the work done by the new economic geographers depends too much on the empirical model and theoretical discussion, and they rarely use their model to analyze real cases in the world. Relative to many other economic geography theories developed in the recent years, the credibility of NEG is unable to compete with them. Similarly, Neary (2001) criticizes the self-reinforcing style of agglomeration process is only a possible result; he called it as the 'propensity to agglomerate'. This is because the analysis of the NEG utilizes too many equation models. However, such criticisms of the NEG seem not to be justified, and in fact, their arguments are too vague and apparently suffer from the bias. For a newly developed economic theory, it is reasonable to base the study on the simplified empirical model and statistical test. Moreover, as discussed earlier, the new economic geographers did apply their models to world case studies. Many scholars have done detailed work on this issue and offered numerous case studies to support the NEG model. In fact, Martin himself also recognizes the case studies made by Krugman.

However, Martin's arguments are not consistent all the time. He seems to deliberately neglect the case studies suggested by the NEG. The case studies of worldwide countries proposed by the new economic geographers are abundant and meaningful. This evidence has generally supported the NEG model. For example, Davis and

Weinstein's study (1999) of Japanese industry location decisions has supported the argument made by the NEG. They found that eight of nineteen manufacturing industries' location decisions were consistent with the economic agglomeration effects. Their empirical model demonstrates that the correlation between the firms' location decision and the agglomeration effect is highly positive. Davis and Weinstein's study suggests that the agglomeration effect plays a key role in shaping the regional production structure of Japanese industrial sector. Moreover, as this chapter discussed earlier, Krugman's study proposes the case of the American manufacturing belt. Krugman and Venables's research (1996) further offers the case studies of American industrial districts to show the striking correlation between industrial concentration and economic agglomeration, such as the automotive centre located in Detroit, the concentration of the garment industries in New York and high-tech industries clustering in Silicon Valley and Route 128 of Boston. In addition, Krugman (1999) points out economic dualism between the developed and underdeveloped regions within Mexico (Mexico City and Chiapas) and Brazil (São Paulo and the Northeast Region).

Furthermore, Martin (1999) criticizes the argument made by the NEG regarding the role of historical incidents in regional agglomeration. This historical argument is flawed and suffers as the "conceptual and explanatory black box" claimed by Martin (1999, p. 76). He claims that the historical factor suggested by Krugman is just metaphorical but not real. Martin (1999, p. 78) alleges that,

The spatial agglomeration models may well predict that, under specific assumptions, industrial localization and specialization will occur, but they are unable to tell us where it actually occurs, or why in particular places and not in others.

Martin's criticism of the NEG is far more sceptical. Although the exact cause of regional agglomeration remains unknown, the new economic geographers have done some detailed works investigating where and why the agglomeration effect actually occurs; how does the agglomeration effect influence regional development via the self-reinforcing process? Besides the historical discussion presented earlier, Fujita and Thisse (1996) give the opinion that history matters for regional convergence and the circular causation tends to make the manufacturing industries concentrate in the same area due to the 'lock-in' effect. They cite the examples of the American industrial belt and European 'blue banana' to support their view. In addition to factors such as trade theory and market location stimulating regional agglomeration, Ottaviano and Puga (1997) suggest that historical accidents can trigger the self-reinforcing feature of agglomeration process.

In addition, Martin's theory suggests that the new economic geographers have deliberately neglected some other important factors which might contribute to the spatial distribution of industries and economic activities, and that factors such as infrastructure construction, institutions and foreign capital might be more important

than the effect of increasing returns to economic scale in shaping regional industrial concentration. According to Martin's argument (1999), the NEG model, with its key characteristics of increasingly economic returns and agglomeration, might be insufficient to explain the complex distribution process of regional industries and other economic activities. However, from my point of view, Martin's analysis has certain limitations on this point. Of course in certain cases, some other factors, such as foreign capital and basic infrastructure, reinforce the effect of regional economic agglomeration and strengthen the location advantage of the existing metropolitan regions; meanwhile, it does not necessarily suggest that the agglomeration effect is insignificant. Moreover, it does not indicate that potential effects of the other factors are stronger than the agglomeration effect. The NEG suggests that agglomeration can add a self-reinforced lock-in effect to the inherited geographical advantage of a region. In other words, the self-reinforcing pattern of agglomeration effect can strengthen the favoured location advantage of one region and make it more attractive to industries and firms, even when its natural geographical advantages become less significant. Therefore, the NEG model provides the key tool to investigate unanswered questions left by the traditional geographic theory, and to analyze the process of regional economic development and unequal distribution of regional industries. Numerous case studies across the world have supported the general analysis made by the NEG. Indeed, this model is important and meaningful, and the NEG should be regarded as the key theory framework in analyzing regional economic development. As Brühlhart (1998, p. 796) highlights, "Industry clustering is a real and significant phenomenon, which cannot be explained as the outcome of a random distribution of discrete plant numbers."

Many criticisms of the NEG model are too vague and inappropriate. Most of the criticisms are not based on the empirical analysis of the world cases. For example, as far as Martin's study is concerned, except for the study of British industries, he did not seem to offer any other credible evidence to support his argument. Therefore, Martin's criticisms are unwarranted and unjustified. Indeed, there are many good reasons to argue that such criticisms are suspect and difficult to match with world reality. At least, this problem has severely limited their power to criticize the NEG.

However, it cannot be denied that the NEG model has its shortcomings. Krugman (1999) recognizes that the NEG might be difficult to put into practice due to crucial equilibrium effects. Issues of spatial industrial distribution and economic agglomeration proposed by the NEG are still short of proper cause investigation, such as the correlation between traffic congestion and regional agglomeration and the precise causes of the self-reinforcing process of agglomeration. As Davis and Weinstein (1998, p. 11) rightly maintain,

Of course, Krugman (1980) cannot be taken straight to data. Such models of economic geography contemplate highly abstract worlds in order to provide clear theoretical insights. Even in such stark models, the inherent complexity of the problems frequently defies analytic solution.

One key limitation of the NEG is this: suppose the model of self-reinforcing pattern of agglomeration effect is right, the self-reinforcing effect should reinforce the economic success and regional industrial concentration in a region, but, this is not the case in some developed regions. Why do some highly developed regions decline and show the tendency of regional economic de-concentration within some countries? Manchester is a good example. Since the British industrial revolution in the mid-eighteenth century, Manchester had become an important economic, trade and industrial centre in Britain. In the past, the regional industrial convergence in Manchester was very evident. However, today, this seems not to be the case. Even if Manchester still remains one the key regional centres within Northern England, its economic power and regional influence have been shaken substantially. This is particularly prominent in manufacturing industry. Equally, this also holds true for Wuhan and Nanjing municipalities in China. The tendencies of economic de-concentration and continually weakening importance in these regions have been demonstrated. What are the causes for this economic decline in some core regions? In addition, if the important role of the historical incidents in affecting regional convergence argued by the new economic geographers (e.g. Krugman) is credible, again, why do some developed core places decline, rather than further develop: in spite of historical advantage? Unfortunately, as far as I understand, these important issues have not been properly addressed and systematically analyzed by the NEG.

Moreover, the new economic geographers still need to provide more empirical studies about the issue of the self-reinforcing feature of agglomeration effect. A lot more analysis and critical discussion of the real cases are still badly needed. In particular, Brühlhart (1998) argues that the case of regional industrial convergence within the EU, claimed by the new economic geographers still needs to be properly tested and further analyzed; this is because the trade and production data of EU has illustrated completely different results on regional industrial agglomeration. Brühlhart seems not to be convinced by the tendency of regional industrial convergence within the EU. Moreover, Kim (1995, cited in Brühlhart, 1998) raises the American case. In America, regional industrial convergence reached its peak around the 1920s, since when, spatial industrial concentration tends to illustrate a decreasing trend; Brühlhart indicates that the strong power of regional agglomeration in the U.S.A has been shaken over the past several decades. Brühlhart (1998) criticizes the ambiguous and indefensible models proposed by the NEG. Furthermore, the NEG model has relied too much on the mathematical equation models and functional forms in understanding regional agglomeration and spatial economic growth. This is precisely what Fujita and Thisse (1996, p. 372) claim,

One of the main limitations of most models of geographical economics is that results seem to heavily depend on strong assumptions made about the economy; in particular very specific functional forms, like the CES or the logit, are used in most models.

4.2.7 Conclusion

The agglomeration impacts on economic growth and spatial income inequality within a nation suggested by the NEG have been assessed. Numerous case studies crossing the world have also been examined. There are some proper conclusions drawn from this section. First, relative to the under-developed periphery, most of the economic components tend to concentrate in the developed core regions due to the effects of high investment returns, scale economies and large home market, such as capital, labour force and technological innovation. In order to enjoy the benefits of increasing returns to economic scale and other externalities generated from the regional convergence of economic factors, the private agents have clustered in a few core regions. Secondly, the developed regions have benefited from the self-reinforcing feature of agglomeration effect; this effect tends to attract the various economic factors relocate to the core regions. Consequently, it is reasonable to claim the agglomeration effect suggested by the new economic geographers plays a crucial role in boosting regional economic growth, and shaping the core-periphery pattern of spatial economic disparity. The historical factors and path dependency are indispensable in understanding the initial stage of regional economic agglomeration. Thirdly, it is certainly the case that the opposite operation forces have brought some challenges to the spatial industrial concentration in the core regions. However, these opposing forces like traffic congestion are very unlikely to halt the tendency toward regional economic convergence. Apparently, the majority of the core regions still maintains a dynamic economy and prosperous society. In general, owing to the economic externalities contributed by the self-reinforcing agglomeration effects, the NEG theory suggests that unbalanced regional economic growth and spatial income divergence are the possible outcome in analyzing regional economy. This is the reason why discussion of NEG should be included in the regional divergence chapter.

The NEG model is a tool to investigate many important questions unanswered by the traditional geography theory. The NEG enables the scholars to understand economic geography and regional economic convergence in depth. However, more empirical investigation and world case studies still needs to be done; this is especially true of the regional industrial cluster and agglomerative economies in China. In addition, the causing mechanisms of self-organizing pattern agglomeration should be further analyzed. In precise terms, what is the initial factor contributing to the starting tendency of regional agglomeration? Moreover, historical accidents alone are not sufficient to explain the complex process of regional economic convergence.

Although the general arguments suggested by new economic geographers are the regional economic divergence and widening spatial income disparity, mainly caused by the scale economies and industrial agglomeration, some new economic geographers have given a quite neoclassical growth suggestion to regional economic development; which is the U-shaped pattern of regional inequality. Krugman and Venables (1995) argue that spatial inequality tends to be narrowed by cheap labour

cost, transportation cost reduction in the periphery and also the impact of economic globalization. Due to the continuous fall of transportation cost in the peripheral region contributed by infrastructure construction, the disadvantage of long distance from the consumers and intermediate-goods suppliers can be largely overcome. Under these circumstances, the manufacturing industries will generate greater incentive to relocate their factories to the under-developed periphery. Krugman and Venables (1995) propose the gradual tendency of regional convergence within some developed countries like America and Japan to support the U-shaped pattern of regional development. In a later study, Krugman (1999) explains his U-shaped regional development in more detail. He states that,

Perhaps more surprisingly, the same model predicts that a continuing decline in transport costs - loosely speaking, the continuing process of globalization - eventually produces a reversal of fortune. ... So there is a second critical point at which industry finds it profitable to move to lower-wage locations. (1999, p. 151)

4.3 Other Theories on Regional Divergence

4.3.1 Myrdal's Theory: Spread and Backwash Effects

Gunnar Myrdal (1957) is perhaps one of the earliest scholars to develop the model of spread and backwash effects for analyzing regional economic growth and spatial income inequality. The spread effects proposed by Myrdal's model are those favorable effects from the developed place (core region) to the under-developed region (periphery), including technology and information spillover, external capital, and advanced managerial skills. By contrast, the backwash effect of Myrdal's model can be regarded as the core regions' economic development generating some negative effects on the economic growth in the periphery. It includes negative effects, such as the capital outflows, skilled human talents and other low-skilled workers from the periphery migrating to the core regions (Hughes and Holland, 1994).

Myrdal's study (1957) suggests the strong correlation between spread and backwash effects, and regional economic development. According to his model, social and economic development in any region would not result in a balanced status between these two effects due to the cumulative causation of social development. Either the spread or the backwash effect will dominate the process of regional economic development; and this unbalanced pattern of development between the spread and backwash effect naturally will lead to the tendency of spatial development disparity between the core and peripheral regions. Due to the strong backwash effects, Myrdal stresses that free market force would tend to make the rich region grow faster and strengthen its economic leading position. Without any governmental policy intervention in free market economy, spatial income disparity would certainly rise. Under the circumstances of free market trade, Myrdal firmly believes that the strong

backwash effect would dominate the process of regional economic development rather than the spread effect.

In particular, Myrdal's study highlights the free trade effect on regional development disparity. Due to the strong effects of increasing returns to economic scale, he points out that free trade would provide more competitive advantages for the rich region which has rich natural endowment and large market potential. Myrdal raises the Italian case to support his argument. With the elimination of internal trade tariffs after Italian unification, the industries located in the Northern provinces have dominated national industrial development, while the Southern provinces remained in slow economic growth. Skinner's study (1977) supports Myrdal's model. In addition, Nicolas Kaldor (1996) has also shared Myrdal's opinion of the significance of free trade on regional economic inequality.

Myrdal's theory of spread and backwash effects is important. On the one hand, Myrdal identifies the effects of trade and market forces in reinforcing spatial economic inequality. On the other hand, more significantly, he has developed the model of spread and backwash effect in influencing the whole process of spatial economic competition between the core and peripheral regions. This model is perhaps one of the most outstanding theories in understanding regional economic growth and spatial inequality. Moreover, Myrdal's model in predicting the future development direction of regional inequality is feasible. According to the arguments made by Myrdal, the degree of spatial economic variation depends on the strength of spread and backwash effects. If the spread effects are stronger than the backwash effects, spatial inequality will be narrowed; otherwise, owing to the strong backwash effects, spatial inequality will inevitably increase and then lead to regional economic polarization. Many scholars have supported Myrdal's model; their studies demonstrate that many countries in the world have been facing the serious challenge of spatial income disparity between the core and peripheral regions; this is because of the absolute dominance of the backwash effects in regional development. The detailed discussion of world evidence will be presented in the next chapter.

Furthermore, Myrdal's study illustrates that regional development and economic prosperity can spread from the core (developed) to peripheral (under-developed) regions during the high level of national development. He indicates that regional convergence would occur when a nation reaches the high stage of economic development. At this stage, the powerful market force will tend to become weak and then the spread effects would dramatically strengthen. Therefore, the tendency toward unequal regional development will cease. North's and Williamson's study (cited in Friedmann and Alonso, 1975) shares this argument that regional economic and industrial development within a nation would converge in the stage of high economic development. Krugman and Venables's research of the NEG (1995) has offered a very similar argument to that of Myrdal. They unveil a so-called U-shaped pattern of regional economic disparity. By suggesting industrial and economic concentration in a few large regions, the new economic geographers also predict that world economic

development will follow the U-shaped feature of change: the process of economic development will first demonstrate economic agglomeration and spatial disparity in the early development period, followed by economic convergence and spatial income equality in “a second stage of convergence in real incomes”, as claimed by Krugman and Venables (1995, p. 859). They further argue that,

It turns out, then, that a relatively simple model predicts a U-shaped pattern of global economic change, of divergence followed by convergence ... (1995, p. 859)

Moreover, in the comparison between Myrdal’s model of spread and backwash effects and the U-shaped regional inequality suggested by the NEG, both studies seem to argue the important effects of the cumulative process of economic factors and increasing returns to economic scale in the developed regions on causing regional economic divergence. Due to these strong backwash effects generated in the core region, the spatial income disparity between the core and peripheral regions tend to increase during the early stage of economic development. However, Myrdal tends to highlight the determining role of free trade and market in contributing to rising regional divergence in this early age. Myrdal suggests that the strong power of free market in regional inequality will decline in the high development age. Consequently, it can lead to the reduction of regional economic inequality. Relative to Myrdal’s model, Krugman and other associated NEG scholars seem to stress that certain crucial factors shape the economic convergence between the core and periphery, such as transportation cost reduction, low wages and economic globalization effect. In addition, in contrast to the new economic geographers, Myrdal shows much more adherence to the idea of state contributing to narrowing regional economic disparity. Myrdal suggests that the government need to actively engage in regional development and do more to help reducing spatial inequality. However, the new economic geographers did not systematically address the institutional issue in achieving balanced regional development. Fujita and Krugman (2004, p. 161) admit that,

In contrast, new economic geography has been concerned with self-organization in space while neglecting developers and governments.

Moreover, Krugman (1999) implicitly indicates that Myrdal’s theory of spread and backwash effects did not systematically consider the important role of economic geography and industrial agglomeration in analyzing the core-periphery pattern of regional disparity. According to Krugman, the NEG attempts to use various empirical estimation and technological models to study regional economy and spatial inequality.

From my point of view, Krugman’s argument of the U-shaped feature of world economic inequality is not completely convincing or reliable. During the development process of a country, the backwash effects seem to be much more powerful than the spread effects in influencing regional economic development, as the evidence presented in the next chapter suggests. The spread effects have great difficulty

generating enough power to reshape the unequal spatial distribution of industries and other economic activities, and then contribute to balanced spatial development between the developed and less developed places. World evidence does not demonstrate this tendency of U-shaped pattern of economic development, and such evidence will be discussed in detail in the next chapter. In fact, Krugman and Venables (1995) themselves claim that this theory might not be able to test the world reality.

Equally, Myrdal's model has certain weakness and shortcomings. Some parts of his argument are problematic. According to the suggestion made by Myrdal, at the high level of economic development, the spread effects will contribute to narrowing regional inequality and then lead to regional convergence. However, the initial condition of the strong spread effects overweighing the backwash effects seems to be very difficult to achieve (if not totally impossible), even if the economic development within a country reaches a high level. The regional economy studied by Friedmann in 1966 clearly suggests that the cumulative process characteristic of industrialization and urbanization tend to favour the economic development in the existing rich core region, while it makes the periphery fall behind and remain in economic backwardness. Friedmann (1966, p. 9) stresses that,

A dualistic structure is thus imprinted upon the space economy, comprising a "center" of rapid, intensive development and a "periphery" whose economy, imperfectly related to this center, is either stagnant or declining.

In fact, Myrdal (1957) himself admits that when the free market forces dominate in the process of national economic development, the spread effects between the countries might be weak, even if the economy reaches the mature development level. In general, when comparing Myrdal's model of spread and backwash effects with the NEG's on U-shaped regional disparity, the important effects of the cumulative process of economic factors and increasing economic scale on regional economic divergence have been highlighted by both studies. Myrdal stresses that the important role of the free trade and market in causing regional economic divergence tends to decline during the later growth period. In addition, Myrdal argues that the role of state in narrowing regional economic disparity should be strengthened. Relative to Myrdal's model, the new economic geographers seem to pay more attention to other factors shaping regional equality, such as transportation cost reduction and low wages.

4.3.2 Friedmann's Theory: Regional Economic Polarization

Friedmann (1966) argues that the firms prefer to remain in the core metropolitan region rather than the periphery; this is because they can take good advantage of rich labour resources, the relatively mature capital market, better infrastructure facilities,

large market and the concentration of other economic factors. In addition, private agents can substantially reduce the production and transportation costs via the effects of scale economies and other external economies. Friedmann's argument should be given some credit. He highlights that the core-periphery structure of economic disparities will persist for long time, once it is established, either at international or interregional level. Friedmann (1966, p. 99) rightly states that,

The spatial structure of these economies is best described by a centre-periphery (spatial disequilibrium) model.

In his later study, Friedmann (1973) points out that the stable core-periphery pattern of regional inequality is due to six key effects generated from the economic growth in the core regions, namely, the dominance, information, psychological, modernization, linkage and production effects²⁴. Relative to the poor peripheral regions, these six major effects tend to reinforce the competitive power of the core region; moreover, they can further stimulate rapid economic growth in the core region. He firmly argues that these six key effects will generate the self-strengthening pattern of regional competitiveness in the core areas; they can enhance the development capabilities of the core regions in dominating the peripheral regions. Moreover, some economic externalities brought about by these six effects will continually attract the private enterprises to agglomerate in a few core places, seeking such benefits and costs saving as technology diffusion and competitive industrial environment.

Friedmann (1973) believes that regional agglomeration can effectively inhibit the decentralization process of modern economies. Agglomeration might be still the strongest power in shaping regional economic concentration rather than de-concentration. In their later study, Friedmann and Wulff (1976) unveil the idea of the core-periphery pattern of regional economic disparity. Friedmann (1992) cites

²⁴ According to the definition given by Friedmann (1973, p. 52-53), "... dominance effect, or the steady weakening of the peripheral economy by a net-transfer of natural, human, and capital resources to the core. Information effect, or the increase in potential interaction within a given core region resulting from its own growth in population, production, and income. Psychological effect, or the creation of conditions favourable to continued innovation at the core, such as rendering the opportunities for innovation more visible, reducing the risks of innovation through imitation, and creating expectations for further innovation. Modernization effect, or the transformation of existing social values, behaviour, and institutions in the direction of greater acceptance of and conformity with rapid cumulative change through innovation. Linkage effects, or the tendency of innovations to breed other innovations by creating new service demands as well as new markets for the services the core region is itself able to supply to other areas. However, not all innovations will have the same capacity in this respect; the multiplier or linkage effects of certain innovations will be greater than for others. Production effects, or the creation of an attractive reward structure for innovative activity operating through the exploitation by innovators of their temporary monopoly position, the appearance of linked systems of innovations, and growing specialization."

cases such as Paris and Tokyo, which are regarded as among the few core economic centres (barely 20 percent of the world's area); however, the majority of regions in the world have largely remained backward in terms of technology innovation and economic growth.

Significantly, Friedmann and Wulff (1976) highlight the weakness of spread effects in contributing to spatial income convergence and economic growth in the peripheral region. They identify the difficulties for the spread effects to emerge and exercise influence in inhibiting the backwash effects. Due to the issue of distance, the disappearance of strong backward linkages in the agriculture and the configuration of economic competitiveness, Friedmann and Wulff (1976) suggest that the spread effect is very unlikely and remains quite minimal.

Thus, it is important to note the following: owing to the strong backwash effects, the trend of economic growth favouring the existing core region rather than the periphery is very difficult to halt. These strong backwash effects dominate the process of regional economic growth and spatial industrial distribution. By contrast, the spread effects seem to be weak and very unlikely to overtake the backwash effects in reshaping the tendency of regional convergence of industries and economic activities.

4.3.3 Hirschman's Theory: the Growth Pole

Hirschman (1959) also indicates that the economic divergence tendency between the peripheral (poor) and core (rich) regions is evident, and will continue to persist for a long time. Hirschman's study suggests that spatial economic disparity is the by-product of economic development within a nation. Whether at international or interregional level, the spatial disparities between the core and periphery, the developed and less developed regions and the prosperous and backward regions will inevitably continue. He believes that the high level of economic development does not necessarily lead to the reduction or even elimination of spatial economic inequality as Myrdal claims. Hirschman criticizes the shortcomings of Myrdal's model. Hirschman (1959, p. 184) stresses that,

In analyzing the process of unbalanced growth, we could always show that an advance at one point sets up pressures, tensions, and compulsions toward growth at subsequent points. But if all of these points fall within the same privileged growth space, the forces that make for transmission of growth from one country, one region, or one group of persons to another will be singularly weak.

According to the key arguments made by Hirschman, relative to the rich North region, the negative backwash effects will inevitably make the economically backward conditions in the poor South region to worsen. The backwash effects will exclude the South region completely from regional economic competition, Hirschman refers to it

as the “several unfavorable or polarization effects” (1959, p. 187-88). For the South region, the few existing industries engaging in manufacturing and other foreign trade-oriented, will shrink and eventually disappear due to the strong competition from counterpart industries located in the North region. Moreover, the ample labour forces and other skilled talents in the South region will increasingly migrate into the North region due to its attractions of high wages and greater working opportunities. In addition, Hirschman (1959) seems to suggest that the externality effects will attract the private enterprises and manufacturing industries to converge in a few core places, such as technology and industrial competition environment, referred to as the “growth pole” by Hirschman (1959, p. 184-85). Consequently, regional industrial agglomeration will emerge and continue; furthermore, the regional divergence between the poor and rich regions will occur.

From Hirschman’s point of view (1959), the firms might ignore the potential economic opportunities and advantages of the under-developed periphery; the over-optimistic estimation of the agglomeration benefits made by the private commercial agents will encourage them to further cluster in the core developed regions which have large market potential and increasing returns to economic scale. Hirschman seems to be firmly convinced by the tendency of regional economic divergence. He argues that regional divergence rather than regional convergence is the natural outcome of economic development. Regional economic growth is necessarily associated with the developed core and under-developed periphery. Wang and Hu’s study (1999, p. 29) also suggest that,

Hirschman simply does not believe that economic progress can appear everywhere at the same time. ... In other words, interregional inequality of growth is a condition of growth itself. Growth, therefore, is necessarily unbalanced.

4.3.4 Endogenous Growth Theory

In general, the so-called ‘endogenous growth theory’ rightly covers the research gap by suggesting technological progress is endogenous rather than exogenous. Romer (1994, p. 3) states, “This work distinguishes itself from neoclassical growth by emphasizing that economic growth is an endogenous outcome of an economic system, not the result of forces that impinge from outside.” In order to measure the contribution of technological progress to economic development, it is essential to take the endogenous growth theory seriously. According to this theory, relative to the neoclassical growth theory, persistent economic development can be achieved by high domestic saving and physical capital investment. The endogenous model has taken knowledge as one kind of capital: under the circumstances, the constant return to capital become possible in understanding the sustainable economic development. Therefore, the slowdown and even stagnation of economic growth due to the diminishing capital return is no longer valid (Mankiw, 2002). Mankiw argues that the

development of technological change and knowledge innovation can be better understood by the endogenous growth theory; whilst Pack (1994) gives this theory credit.

Charles Jones is respected for his analysis of the impacts of endogenous technological diffusion on long-term regional economic growth. Jones (1995) suggests an example of endogenous economic development which is generated from R&D investment activities and based on technological progress. In order to develop new technology and enhance the capability of technology innovation, in his later study, Jones indicates that the state and the other private agents should spend much more money on R&D-based research activities. The crucial role of technology in stimulating regional economic development should not be underestimated. Jones (1999) seems to maintain that the only possible way to achieve increasing return to economic scale must be based on the new innovation idea and technology adoption. As Jones (1999, p. 143) points out,

However, the property those ideas are nonrivalrous means that growth and increasing returns to scale are tightly linked.

However, endogenous growth theory is not without criticism. Smith et al, (1993, cited in Martin and Sunley, 1998) suggests that the over-rigidity of endogenous growth inappropriately eliminates the technological contribution made by the small enterprises. They believe that the endogenous model is unjustified in claiming that technology innovation is oligopoly. In addition, the research by Martin and Sunley (1998) suggests the limitations of the endogenous growth theory. They argue that one of the important limitations of the endogenous theory is its weak explanation regarding the development functioning of technology spillover among the different regions. The endogenous growth theory has failed to address the difficulty of regional income convergence in the real world.

The research conducted by Martin and Sunley (1998) has also made criticisms of the endogenous growth model. In particular, first, they claim that the lack of empirical evidence of this model weakens its ability to explain the formation mechanism of increasing returns to capital in the individual industries. The endogenous theory over depends on mathematical assumption and statistical exercises. Secondly, Martin and Sunley claim that regional differences and distinguished geographical factors have not been properly considered in the study of growth mechanisms proposed by the endogenous growth model. In general, they are not fully convinced by the constant capital returns suggested by this model.

The endogenous growth theory enlarges our knowledge about potential sources of long-term economic development. The technological factor is the key to sustainable economic growth. The pace of regional development depends on the speed of technology diffusion. Moreover, R&D investment made by the state has played a determining role in importing and imitating the advanced technology. McCallum

(1996) suggests that the endogenous growth theory can properly explain long-term economic development and varying rates of economic growth among the different nations; and it can estimate the impacts of governmental policies and new technology on changing the equilibrium ratio of capital in labour.

4.4 Final Conclusion

In the first section of this chapter, the main theoretical framework proposed by the NEG has been discussed. The NEG model is an important approach to analyzing many important questions unaddressed by the traditional economic geography. In general, there are some conclusion which can be summarized. First, relative to the economically poor regions, the majority of economic factors such as capital, labour and technology tend to converge in the developed regions mainly due to the effects of economic externalities and large market. In seeking the benefits of increasing returns to economic scale and other externalities generated from regional economic convergence, the various industries and firms have clustered in a few developed core regions. The convergence of economic factors is an important factor causing regional industrial cluster and agglomerative economies. Secondly, the core regions have benefited substantially from the self-reinforcing feature of agglomeration effects; these effects tend to further attract the economic components to relocate to these developed regions. The strong agglomeration effects contributed by regional economic convergence reinforce the inherited geographical advantage of the prosperous coastal areas. Therefore, it is reasonable to claim that the agglomeration effect suggested by the NEG plays a crucial role in affecting regional development, and shaping the core-periphery pattern of spatial economic disparity. Moreover, history and path dependency are important in understanding the initial stage of regional economic agglomeration.

Thirdly, opposing forces have brought serious challenges to the regional industrial cluster and economic agglomeration in the developed core regions; this is particularly due to heavy traffic congestion. However, these forces have difficulty in halting the development tendency toward regional convergence. These opposition factors seem to be too weak to generate sufficient power to change industrial clustering. As the world evidence demonstrates, the majority of the existing developed regions still maintain their prosperous economy and relatively fast growth of industries and other economic activities. In conclusion, due to the self-reinforcing effects of agglomerative economies, the new economic geographers claim that unequal regional growth and spatial income divergence are the two likely outcomes in understanding long-term regional economic development.

However, this NEG model faces its own serious problems. More empirical analysis and world case studies need to be done. Moreover, the forming mechanisms of self-organizing pattern agglomeration should be further investigated. Neither was I fully convinced on the rôle of history in causing regional agglomeration. If the

economic agglomeration theory suggested by the NEG is justified, the self-reinforcing effect should reinforce economic development and regional industrial convergence in an existing developed region; however, this is simply not the case in some regions. Why did some previously prosperous core regions decline?

In the second part of this chapter, other main theories of regional economic divergence have been briefly discussed, including Myrdal's theory on spread and backwash effects, Friedmann's theory of regional polarization, Hirschman's growth pole and the endogenous growth theory. In general, although there are various distinctions between the theories, theoretical framework and empirical model proposed by these scholars, overall, they suggest that regional economic development show the tendency of regional income disparity and economic divergence rather than spatial economic equality and convergence. In particular, Myrdal and Friedmann's theories have been emphasized by this study.

My general arguments are that the theory of spread effects outweighing the backwash effects during the high level of economic development claimed by Myrdal, is highly problematic. The real power of spread effects in influencing regional convergence is perhaps largely determined by regional economic agglomeration and geography. The strong backwash effects in shaping the spatial disparity between the core and peripheral regions are very difficult to inhibit. The powerful backwash effects dominating regional development and spatial industrial distribution are notable. The spread effects seem to be weak and unlikely to overtake the backwash effects in contributions to regional economic convergence. In addition, Hirschman's growth pole and Friedmann's core-periphery models deserve to be taken seriously. Both Hirschman and Friedmann highlight the tendency of regional economic divergence, but not convergence. Numerous case studies range from the developing to the developed countries will be assessed in the next chapter to demonstrate that a high level of economic development has not automatically led to the reduction of spatial economic disparities suggested by Kuznets, and other neoclassical growth scholars. The tendency of regional economic convergence has not occurred to any desirable extent.

According to the NEG, self-reinforcing industrial cluster and agglomeration economies are the crucial factors explaining the concentration of economic factors and regional economic growth. As a result, the core-periphery pattern of disparity and unbalanced regional development are the likely outcomes. Moreover, the previous research produced by Myrdal, Friedmann and Hirschman all suggests that economic development is inevitably correlated with widening regional disparity and economic divergence, but not with spatial equality and economic convergence. Based on the discussion of these theories, I am in an appropriate position to discover and explain the regional development phenomenon in the context of Guangdong. If these theories are credible, the development tendency of regional inequality in Guangdong should also show a divergent trend. Industrial clustering might be an important cause of regional economic growth within Guangdong. The prosperous PRDA might have

benefited from industrial cluster and economic agglomeration during the reform period. Detailed discussion of the regional inequality tendency and industrial clustering in Guangdong is presented in Chapters Nine and Eleven.

Chapter 5 Literature Review Part IV: Regional Economic Convergence

5.1 Introduction

Before analyzing the changing trend of regional disparity in Guangdong, the literature regarding regional convergence theories also needs to be discussed with reference to world evidence. This chapter attempts to provide a critical discussion of the theoretical literature on regional economic convergence theories. The chapter attempts to produce the discussion of the literature on regional convergence theories. The first main section briefly discusses Kuznets's theory and estimates the importance of his inverted U-curve model in analyzing regional economic development and spatial income disparity. The neoclassical growth theory will be assessed in this section. The second main section presents the world evidence discussion on the developing trend of regional economic inequality.

The specific structure of this chapter is as follows. Section Two emphasizes discussion of the main theories of regional income convergence. In this section, the inverted U-curve theory of regional inequality proposed by Kuznets is analyzed. The neoclassical growth theory is also discussed in this section. In particular, Barro and Sala-i-Martin's studies have been highlighted. Section Three aims to present detailed evidence review of evidence regarding the tendency of spatial economic development. The specific case studies are assessed in order to properly estimate the credibility of the main regional convergence theories in explaining the real world. Section four draws final conclusions.

5.2 Regional Economic Convergence

5.2.1 Kuznets's Hypothesis: the Inverted U-curve of Regional Inequality

Likewise, the various factors that have been suggested above would explain stability and narrowing in income inequality in the later rather than in the earlier phases of industrialization and urbanization. (Kuznets, 1955, p. 18)

In early research, Kuznets (1955) unveiled his classic model of the inverted U-shaped of regional inequality. To some extent, Kuznets's influential theory is similar to Myrdal's model. Due to the inherited characteristics of economic development and technological innovation, Kuznets (1953) argued that high-speed economic growth as demonstrated in increasing per capita income cannot occur in all countries at the same time. Relative to the developed countries, the economically poor countries have faced

various challenges and difficulties in adopting and imitating advanced production methods and new economic mechanisms. Therefore, the rising regional income disparity between the rich and poor countries is not so surprising. Economic development is inevitably linked to unbalanced growth and rising income inequality among the different nations.

In his later study, Kuznets (1974) found that there were huge gaps in personal income and economic growth between the highly developed (e.g. America and West Europe) and under-developed countries (e.g. Sub-Saharan Africa and Asia), and also between the Communist (e.g. China) and non-Communist countries (e.g. Japan). More significantly, this huge variation of regional income across the world has dramatically widened over the past two centuries. It is important to note the following points. Kuznets (1971; 1974) recognizes that the poor countries have recorded the dramatic rate of economic growth. However, relative to the equivalent figure in the under-developed countries, Kuznets identified that the average growth rate in the developed countries was still far higher. According to the theory proposed by Kuznets, national economic development follows some specific rule: first leader, then follower and lastly, the later catcher. As Kuznets (1966, p. 468) pointed out,

The model of spread of an economic epoch thus suggested inevitably produces differential rates of growth among the pioneers, early followers, and late followers ...

Kuznets (1955; 1966) presents detailed causal explanation regarding the inverted U-curve of regional inequality. Both the high birth and low death rates in the economically poor countries had negative effects on their economic development. Moreover, during the early stage of economic development, some favourable factors (e.g. the state-oriented policies and physical geography) and advanced industries in the developed countries will reinforce their leading position of material wealth and further enhance their capabilities in pursuing the high-speed economic growth. If the political stability can be guaranteed and institutional failures can be prevented, Kuznets (1966) seems to indicate that the fast economic growth in the under-developed countries would be possible. In general, he argues that economic growth can be achieved in all countries; regional income convergence will occur in the long term. In the high development stage, inevitably, regional economic inequality would be narrowed.

In the search for factors causing regional income divergence and the varying rates of economic growth among the nations, on the one hand, Kuznets (1953) points out that the industrial system should be blamed for this divergent tendency. Slow spread of the industrial system²⁵ between different countries naturally leads to regional inequality.

²⁵ Kuznets offers his explanation regarding the concept and characteristics of the industrial system. Kuznets (1953, p. 238-39) maintains that, "... the 'industrial system' – a concept used to designate a wide application of knowledge, based on empirical science, to the problem of economic and

Kuznets (1966) argues that the impact of inherited geographical location on spatial income divergence should not be underestimated. He cites evidence from nations across Europe to support his argument. Only the countries and regions located adjacent to the Atlantic Ocean, such as France and Great Britain have had access to trade and the capitalist economy since the period of the 15th century. However, many Italian and German regions, and the Central and Eastern European countries, which have relatively poor location, had difficulty taking advantage of trade and industrial expansion during that time, Kuznets (1966, p. 466) refers to it as “the epoch of merchant capitalism”. Kuznets further stressed that,

These arguments suggest that in any breakthrough to a new epoch, only a few countries can satisfy the requirements of successful pioneering ... (1966, p. 466)

On the other hand, Kuznets (1966) claims that political factors in the under-developed countries like poor state policies and a lack of government administrative experience, have also played an important role in causing the regional economic disparities and widening income gaps between the developed and underdeveloped countries following political independence.

In my opinion, to some extent, Kuznets’s model of U-curved regional inequality is attractive. It indicates that the importance of geographical factors in widening regional income inequality should not be undervalued or even neglected. Geography makes a difference to regional economic development. Nielsen and Alderson (1997) give their support to the Kuznets model. More significantly, Kuznets’s study suggests that the tendency toward regional income convergence is a long journey. The way leading to regional convergence is beset with difficulties (supposing the hypothesis of U-shaped regional inequality is not a null one). In addition, Kuznets’s research identifies that the state can be a negative factor in influencing regional economic development and economic convergence rather than purely a contributor to regional balanced development as many people think, for example, in cases of state failure and poor policymaking.

However, Kuznets’ U-curve model suffers from numerous serious problems and has faced strong criticism from academics. How credible is the U-shaped hypothesis of regional inequality? How useful is Kuznets’s model to real world analysis of regional economic development? These questions need to be properly addressed. I am not convinced by the key argument made by Kuznets, which is the tendency of regional income convergence in the long run. Kuznets’s answers to some key questions are not satisfactory. In fact, it is problematic. For example, why does regional income inequality tend to decline at the high development stage? What is the main cause contributing to spatial disparity reduction during the high development period? If the scale return of technological diffusion is increasing, can the developed countries

social technology.”

maintain their high-speed economic growth during the high development stage? As far as the under-developed countries are concerned, what is the effective way to apply advanced production methods and new economic mechanisms during the economic development process? In addition, some important issues have been left entirely unaddressed by Kuznets. Relative to the high development stage, why are the factors of demographic and other state-oriented policies so powerful in causing regional divergence in the early development stage? What is the determining force contributing to the faster economic growth in the poor counties than the rich ones in the high stage of economic growth?

Due to the high salaries of skilled workers, Aghion and Howitt (1998) find that regional income inequality might further increase: according to their studies in the developed countries. They argue that the trend of regional inequality can be reflected in a cyclical rather than inverted U-shaped pattern. The empirical research conducted by Anand and Kanbur (1993a; 1993b) has demonstrated the opposite trend of regional inequality development rather than the U-curved hypothesis proposed by Kuznets. The general level of regional inequality has not changed during the whole period of economic development (Anand and Kanbur, 1993a). They state that,

We showed that under the Kuznets assumptions the entire Lorenz class of inequality indices increases at the start of the process, but the behaviour of inequality at the end of the process - and hence the existence of a turning point in the inequality development relation - is ambiguous. (1993a, p. 35-36)

5.2.2 Neoclassical Growth Theory

Solow (1956; 1967) is one of the scholars renowned for his early contribution to neoclassical growth theory, whilst others have developed the theory further. According to the neoclassical growth model, in order to achieve economic growth, capital investment mainly based on the high rate of domestic savings, is crucial. In addition, the low growth rate of population is essential. This is because high population growth will cause the per capita investment of the workers to become less and insignificant. High population growth in a region inevitably leads to insufficient capital investment which then negatively affects economic growth. (Whiteley, 2000)

Moreover, due the effect of decreasing returns to capital scale, the neoclassical growth model suggests that a regional economy solely depending on capital investment will cease to grow in the long term. Capital is not the sustainable way to stimulating long-term economic development (Aghion and Howitt, 1998). Barro and Sala-i-Martin are two influential researchers of neoclassical theory. Their studies demonstrate the tendency of regional economic convergence. Everything else being equal, if the per capita income in the poor regions is lower than the average steady-state level of income, Barro and Sala-i-Martin argue that these regions tend to grow faster than the rich regions, and regional economic convergence will occur in

the future. In terms of neoclassical language, this is because of the diminishing returns to capital scale. Moreover, in his early study, Barro seems to argue that reduction of the economic inequality between the poor and rich countries needs to be accompanied by sufficient human resources. Barro (1991) identifies the importance of human capital in achieving rapid economic growth in the under-developed regions. However, the research conducted by Barro and Sala-i-Martin (2004) further points out that regional economic convergence is not unconditional but conditional; it depends on the population growth rate, human resources and state-oriented policies.

In fact, conditional regional convergence²⁶ claimed by Barro and Sala-i-Martin indicates that spatial income equality does not necessarily occur. If the under-developed peripheral regions have insufficient human capital and technological innovation, regional economic divergence might continue or even worsen. The studies of the neoclassical growth model conducted by Barro and Sala-i-Martin and endogenous growth theory are both important. Although there are clear distinctions between these two theories on the original cause and development of technological progress, both studies suggest that technology matters for economic growth in the poor regions.

Using neoclassical empirical estimation, Barro and Sala-i-Martin (1992) have referred to case studies from America, Japan and other developed countries to support their argument of regional economic convergence. For example, in terms of income per capita, they found that the poor states in the USA had grown at a faster rate than the rich ones during the period since the mid 19th century. In their later study, Barro and Sala-i-Martin (2004, p. 496) further point out that,

The results indicate that absolute β convergence is the norm for these regional economies. That is, poor regions of these countries tend to grow faster per capita than rich ones.

In general, Sala-i-Martin (1996) concludes that the pace of economic convergence between the poor and rich regions across the world is around two percent per year by expanding his empirical evidence to the OECD countries. According to the research conducted by Sala-i-Martin (1996), the slow-speed tendency toward regional convergence is due to slow technological spillover, and the high cost of technological imitation across the different countries. In terms of the neoclassical model, for the under-developed regions, Barro and Sala-i-Martin (1997) argue that the only key and sustainable source for achieving fast economic growth in the long run is the

²⁶ Aghion and Howitt (1998) point out the definition of conditional convergence in comparison to absolute convergence. They state that, "Absolute convergence takes place when poorer areas grow faster than richer ones whatever their respective characteristics, whereas there is conditional convergence when a country (or a region) grows faster the farther it is below its own steady state. The latter form of convergence is definitely the weaker. Under certain conditions, conditional convergence even allows for rich countries to grow faster than poorer ones" (1998, p. 31).

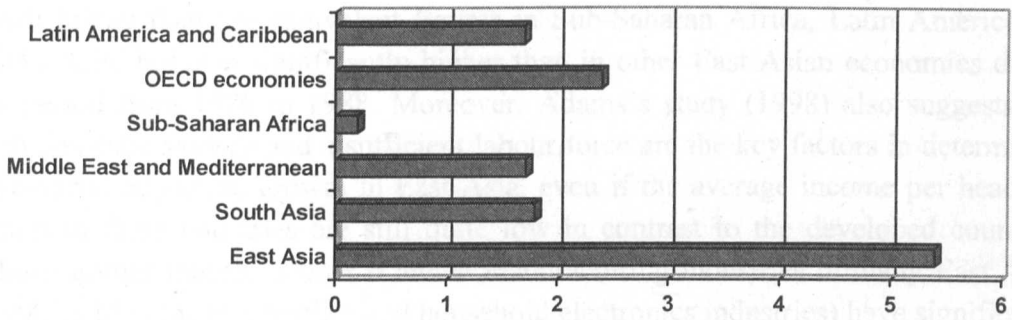
technological factor; they refer to it as the “exogenous technological progress” (1997, p. 1). Both the physical capital and labour force are critical only in the short-term period. Furthermore, Barro and Sala-i-Martin highlight that regional economic convergence must be based on the model of endogenous technology-oriented growth. In order to achieve the high-speed economic development, they indicate that the economically poor countries and regions need to focus on technological progress and spend more public investment on R&D activities. In their later research, once again, Barro and Sala-i-Martin (2004) stress the importance of technological factors in contributing to long-term economic development.

According to the neoclassical growth theory, when reaching the steady state level (equilibrium), it is only technological progress which would shape the output growth per worker rather than the high growth of capital and labor force due to the factor of diminishing capital returns (Mankiw, 2002). In other words, technological progress matters most. However, Mankiw (2002) claims that one of the main problems regarding the failure of Solow’s neoclassical model is in explaining the origin and causal mechanism of the technological progress. Neoclassical economists simply suggest that the technological progress is exogenous.

The neoclassical growth theory has gained support from certain scholars. Chen and Fleisher’s study of China has provided evidence to support the credibility of neoclassical growth theory. As far as economic growth rate and per capita income are concerned, Chen and Fleisher (1996) maintain that there is a strong tendency toward the conditional convergence of regional income at the provincial level. In addition, Li, Liu and Rebelo (1998) also cite the case of China to support the neoclassical model. Based on the empirical estimation of provincial data in China between 1978 and 1995, those Chinese provinces with a sufficient labour force and low population growth have achieved higher rates of economic growth than their counterparts with insufficient human capital and high population growth.

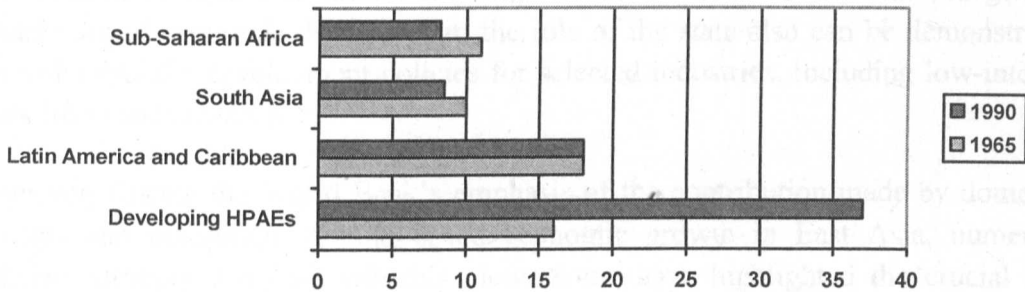
The case of East Asian economic success has been argued by the neoclassical economists and World Bank to support the neoclassical growth theory. The general argument made by the neoclassical growth model suggests that the East Asian economic miracle lies in the high growth rate of domestic savings, capital investment and labour force reserves. Their economic development highly depends on the export-oriented trade, high growth rate of domestic savings; abundant skilled labor forces (see Figs 5.2 and 5.3 below for more details). In a highly influential book, ‘*The East Asian Miracle*’, the World Bank (1993) suggests the crucial roles played by domestic investment and labour force in the rapid economic growth in East Asia.

Fig 5.1: Average Growth Rate of World per Capita GNP, 1965-90



Reproduced from World Bank (1993, pp. 2)

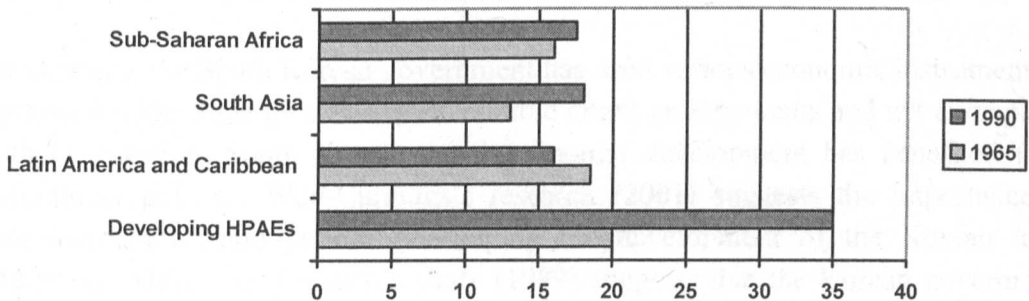
Fig 5.2: Average Percentage of Savings in GDP



Note: The regional averages are un-weighted.

Reproduced from World Bank (1993, pp. 41)

Fig 5.3: Average Percentage of Investment in GDP



Note: The regional averages are un-weighted.

Reproduced from World Bank (1993, pp. 41)

Moreover, as far as East Asia is concerned, the neoclassical economists argue that the economic success of this region has benefited from dynamic trade development. For example, according to the data given by the World Bank (1993), Singapore's economic miracle is mainly due to export-oriented free trade. In terms of the ratio of

total trade to GDP, the equivalent figures in Singapore were indeed outstanding between 1970 and 1988. The ratios of total trade to GDP in Singapore were not only much higher than the equivalent figures in Sub-Saharan Africa, Latin America and South Asia, but also significantly higher than in other East Asian economies during the period from 1970 to 1988. Moreover, Adams's study (1998) also suggests that high domestic savings and a sufficient labour force are the key factors in determining high-speed economic growth in East Asia, even if the average income per head and wages in these countries are still quite low in contrast to the developed countries. Adams argues that the labour-oriented manufacturing industries in many East Asian countries (e.g. the toy, textiles and household electronics industries) have significantly benefited from their rich labour resource. Moreover, Yusuf (2001) and Stiglitz (2001) have both suggested the importance of domestic savings and investment in boosting the economic success in this region. The World Bank (1993) does point out that although these are determining factors, the contribution made by the government in influencing economic growth should not be ignored. It suggests that the state intervention in economic affairs has emphasized the rule of law and the general framework of economic development; the role of the state also can be demonstrated from the specific development policies for selected industries, including low-interest bank loans and subsidies.

However, despite the World Bank's emphasis of the contribution made by domestic savings and investment to high-speed economic growth in East Asia, numerous scholars strongly disagree with this view. Some have highlighted the crucial role played by the institutions in guiding industrial development and stimulating local economic growth in East Asia. A strong state is vital for the economic growth miracle in this region. Based on the various economic and non-economic channels, such as tax, exchange rate and other administrative methods, Yusuf (2001) argues that the government intervention and state-oriented development policies have played an important role in achieving industrialization and modernization in East Asia.

For example, the South Korean government has used various economic instruments to promote foreign trade growth like favourable credit arrangements and tax exemption. Without question, South Korean rapid economic development has benefited from institutional policies. Woo-Cumings's research (2001) suggests the importance of state intervention and policies in helping the development of the Korean large enterprises. Moreover, Amsden's study (1989) suggests that the Korean government has supported economic development and investment growth through the channel of price distortion. In particular, Amsden stresses that the state, rather than the market, has played a crucial role in industrial investment decisions, resource allocation and the development of private enterprises.

Both the studies conducted by Amsden (1989) and Chang (1993) highlight that the influential role played by the Korean government in stimulating industrial and economic development lies in the state capability in disciplining the private

enterprises²⁷. Chang (1993) further highlights the importance of the government and state-oriented industrial policies in shaping the rapid development and expansion of selected industries. The Korean government adapts different channels to support the growth of certain industries, such as special investment funds and tax treatment, subsidies and preferential credits; these financial supports are regarded as the 'state-created rents' by Chang. In addition, industry entry restriction and the protection of import products and tariffs have been utilized to restrict industrial competition and encourage the development of domestic industries; called the 'designated industries' by Chang. In general, he summarizes the major contribution made by the state in stimulating the industrial development of South Korea,

A constant upgrading of the industrial structure based on the development of local technological and managerial capabilities was seen by Korean policy-makers as the surest way to achieve sustained growth and efficient structural change, and hence higher living standards. (1993, p. 153-54)

As far as Taiwan (China) is concerned, Robert Wade's study (1990) suggests that the dramatic economic development on this island region since World War II has benefited from the political stability and authoritarian one-party regime. Moreover, high domestic savings were also vital for the rapid economic growth in Taiwan (China) during the period between 1955 and 1985. By adopting various instruments to discourage private house ownership, the government has played an important role in contributing to the high growth rate of domestic savings (Wade, 1990). In general, Wade (1990) highlights that the neoclassical economists have ignored the significant role played by the state in controlling market and shaping the industrial development direction of Taiwan. Various industrial policies and controls have been employed by the government to stimulate industrial growth and strengthen the international competition capabilities of the domestic industries.

In fact, the fast industrial development and economic success in Japan during the high-speed development period has benefited substantially from the strong state. In studying Japanese economic development during the post-World War II period, the institutional factor simply cannot be underestimated. Chalmers Johnson (1982) conducts some detailed research of this issue. He believes that the Japanese government and its administrative guidance was key to the dramatic economic development of this nation²⁸. In particular, Johnson points out that the various industrial policies issued by the Ministry of International Trade and Industry (MITI) have successfully contributed to the fast industrial expansion and economic growth in Japan. A later book '*Japan: Who Governs? The Rise of the Developmental State*',

²⁷ For example, Chang (1993, p. 150) argues that, "The Korean state played a central role in the country's economic development. Of course, such a result was only possible because the Korean state was a strong state which could discipline firms."

²⁸ To be fair, Johnson (1982) discusses the weakness and limitations of the state in guiding the market and economic activities of Japan.

Johnson (1995) suggests the idea of economic nationalism²⁹ and capitalist developmental state³⁰, which are the two key principles used to analyze state influence on Japanese economic development during the 1960s. Based on these two principles, various government strategies were designated and implemented to stimulate industrial expansion and high-speed economic growth, such as the policies of industrial development and foreign aid.

However, the role of the strong state in boosting East Asian development is arguable. With reference to the World Bank's opinion, the argument of the strong state contributing to the East Asian economic miracle has been criticized by some scholars. For example, Adams and Vernon (1998) pointed out that the economic policy failures which were caused by the state, had indirectly led to the 1997 East Asian financial crisis. This is especially true of inappropriate financial and exchange rate policies. In addition, the open and uncontrolled financial markets in East Asia, without tight government supervision, should be blamed for the economic crisis and economic recession in these countries since 1997. Adams and Vernon (1998, p. 205) argue that,

It is apparent, then, that fiscal and financial policies played a role in the crisis. Earlier application of appropriate policy remedies might have eased the crisis.

To be fair, Woo-Cumings (2001) also recognizes the limitation and problems of the government intervention and state-oriented policies which accounted for the South Korean economic crisis. Woo-Cumings has made detailed discussion regarding the state's failures. In particular, she highlights the failures of the state-oriented policies implemented by the Kim Dae Jung administration. Woo-Cumings (2001, p. 358) points out that, "The cause of the crisis has to be sought in the failure of the reform efforts in Korea."

My general argument is this: to some extent, the arguments made by the neoclassical economists are reasonable. It is justified to claim that domestic savings, foreign trade and capital investment have contributed to East Asian economic success. The dramatic trade development, high growth rates of investment and savings are certainly impressive, according to worldwide comparisons. However, more significantly, the strong role played by the state in affecting the economic development of East Asia

²⁹ "By economic nationalism, then, I mean that the Japanese pursue economic activities primarily in order to achieve independence from and leverage over potential adversaries rather than to achieve consumer utility, private wealth, mutually beneficial exchange, or any other objective posited by economic determinists." (Johnson, 1995, p. 105-06)

³⁰ The concept given by Manuel Castells (cited in Johnson, 1995, p. 67) regarding the capitalist developmental state is this, "A state is developmental when it establishes as its principle of legitimacy its ability to promote and sustain development, understanding by development the combination of steady high rates of economic growth and structural change in the productive system, both domestically and in relationship to the international economy."

should not be overlooked. The institutional factor accounts for the so-called East Asian economic miracle. The regional development experience in East Asia indicates that the government has heavily involved itself in the development of industries and other economic activities by adopting various state-oriented preferential policies. South Korea (e.g. discipline of private enterprises by the state), Japan (e.g. capitalist developmental state) and China (Special Economic Zones) are three classic examples. As far as this point is concerned, some parts of the view suggested by both the neoclassical growth theory and the World Bank are inappropriate and problematic. They seem to underestimate the important role played by the state in stimulating the economic growth and industrial expansion in East Asia. It is clear that domestic savings, trade and labour force are not the whole story in explaining the economic success in this region.

5.3 World Evidence Discussion: Regional Inequality Tendency

5.3.1 Evidence against Kuznets's Hypothesis

In general, the evidence of spatial economic development across the world does not support the inverted U-curved hypothesis and regional income convergence of Kuznets and the neoclassical economists. The trend of spatial income convergence has not occurred to any desirable extent. By contrast, the development tendency of regional divergence is notable. For example, Rati Ram strongly criticizes Kuznets's model by adopting the regression estimation of countries throughout the world. Instead of the inverted U-curve trend of regional inequality, Ram (1991) points out that the uninverted U-shaped model is more suitable for the real world situation. Ram (1991, p. 1113) maintains that,

On the contrary, the regression estimates fit well on uninverted U-curve, indicating that inequality first declines with economic growth and then rises after reaching a bottom.

Ram points out the case of the USA to support his uninverted U-shaped model. Based on his time-series and cross-sectional estimation of the American states, Ram (1991) did not find the predicted result of inverted income inequality hypothesis suggested by Kuznets. In his later study, Ram (1997) applies his uninverted U-shape model to some other countries in the world. The regression test virtually confirms his argument rather than Kuznets's model. Many countries have experienced the tendency of first decreasing (in the 1950s and 1960s) and then dramatic rise (starting around the 1970s) of regional disparity. Ram's research (1997) has assessed comprehensive data from nearly all of the economically rich and rapidly developed countries. Significantly, Ram found that the regional development experience of the majority of countries is very similar to America, and these cases had rejected Kuznets' inverted U-curve hypothesis.

Moreover, based on their consistent data selected from the sample size of 60 nations, Anand and Kanbur (1993b) point out that the development tendency of regional inequality has contradicted the popular U-shaped inequality hypothesis. Moreover, Bowman's study (1997) has strongly questioned Kuznets's inverted U-shaped pattern of regional inequality, and his case studies seem to completely reject Kuznets's argument. Costa Rica has experienced relatively fair and equal economic growth among the various regions; Bowman (1997, p.134) refers to it as the "case of growth-with-equality". Greece is another good case of economic development accompanying regional equality. As far as Asia is concerned, the evidence is more prominent. For example, Japan has not experienced the inverted U-curve trend of economic regional inequality. In general, Japan escaped serious regional inequality during its period of high-speed economic growth. Moreover, South Korea's development experience also completely rejects Kuznets's claim. Taiwan (China) is another widely discussed example. It has illustrated an impressive trend of relatively fair economic development and regional income equality during the post-World War II period. (Bowman, 1997)

The evidence from numerous countries seems to suggest that regional income inequality does not follow the inverted U-shaped hypothesis. Rising regional inequality does not necessarily occur in the early development stage alone; regional economic gap might also increase in the later period. Equally, the regional income divergence between the rich and poor regions will not inevitably decline during the high development stage. World evidence presents a very complex and unpredictable trend of spatial income inequality. Kuznets's argument and his U-shaped regional disparity model seem to contain flaws. His model might be largely based on the personal speculation and subjective claims. As he honestly admits,

... I am acutely conscious of the meagerness of reliable information presented. The paper is perhaps 5 percent empirical information and 95 percent speculation, some of it possibly tainted by wishful thinking. (1955, p. 26)

5.3.2 Evidence against Myrdal's Spread Effect Thought

Myrdal's theory of the strong spread effect stimulating economic growth in the under-developed nations, and aiding reduction of spatial inequality, is supported by Puga and Venables (1996) in their case study of East Asian countries. They firmly allege that Japanese economic development has successfully spread into other East Asian countries. Puga and Venables (1996) argue that the high costs of manufacturing production and labour have caused the private firms which had clustered in the core region to relocate to peripheral regions; consequently, this triggered the industrial dispersion process. Therefore, the core-periphery pattern of regional economic disparity will tend to be narrowed. As they state,

Industrialization then commences in this country, and takes place at a

rapid rate as forward and backward linkages are created and a critical mass of industry attained. The process may then repeat itself, so industrialization takes the form of a sequence of waves, with industry spreading from country to country. (1996, p.460)

To my understanding, citing the East Asian case to illustrate the strong spread effect claimed by Myrdal, Puga and Venables's view is completely inappropriate and problematic: it cannot be justified. Their argument regarding the Japanese economic development spreading into other East Asian countries is hardly true. Except for South Korea, Singapore and a few small inland economies like Taiwan (China) and Hong Kong (China), many East Asian countries are still poor and economically backward in terms of absolute income levels (in contrast to Japan). More significantly, some East Asian countries (see Fig 5.1 below) have not seemed to benefit significantly from the spread effect of the Japanese high-speed economic growth during the post-World War II period. As discussed earlier, the rapid economic success in East Asia is mainly due to the high growth of domestic savings and capital investment, and the strong state. Of course, Japanese investment has played an indispensable role in the industrial development of some East Asian countries, such as the automobile industries in Thailand and Malaysia. However, the fundamental factors causing the East Asian economic success lie in high domestic savings, rich labour reserves, free trade and also government institutions, but not to the Japanese factor.

Furthermore, Henderson, Shalizi and Venables (2001) stress that there is no comprehensive and conclusive world evidence to support the trend of regional convergence. More significantly, no conclusive evidence can demonstrate in what situation, when, and under what kind of government policy, regional convergence will occur. The development experience of many developed regions and countries clearly illustrates that crossing the threshold from strong backwash effects to strong spread effects did not come about to any significant extent. In fact, some previous researches claim that Myrdal's argument is unreliable and problematic. These studies demonstrate the continued existence of regional disparities within both developing and developed countries.

For example, the European Union is a highly developed region where significant regional variance still exists. Combes and Overman (2003) found that the regional disparity (as shown in per capita income) exists in this Union. They argue that, "the map clearly demonstrates the strong core periphery pattern which sees rich regions located on a 'blue banana' running from the South East of the UK through Holland, West Germany and then curving round (hence the banana) through Austria and into Northern Italy" (2003, p. 8). Moreover, in Italy, which is a developed and economically rich country, significant regional economic disparities have been found. The northern areas of Italy are generally much more developed and prosperous than their southern counterparts. Kaldor (1996) argues that the industrial productivity in the southern region has been around 20 percent lower than in the north. According to

his Italian case study, he firmly believes that the spread effort has never become strong enough to outweigh the backwash effect. Moreover, Friedmann (1966) offers his study on Western Europe and America. He argues that both sides of the Atlantic still suffer from the serious problem of spatial income inequality between the rich and poor regions. America and Western Europe, which are economically rich regions, both have reached high level of economic development and achieved national industrialization; however, the enormous regional gap between the developed and under-developed areas is notable.

In conclusion, the outweighing of backwash by spread effect at high levels of economic development, suggested by Myrdal, has never come about to any desirable extent. In my view, the power of the spread effect in reshaping regional development is perhaps largely determined by regional agglomeration and precise geographical location rather than the economic developmental level. It seems: the further away from the core region, the smaller are agglomeration affect, the weaker the spread effect. Therefore, the relatively long distance between the peripheral and core region tends to keep the periphery in poverty, and reinforce its backwardness and economic stagnation. For example, in a case study of America, Hanson (2001) found that the economic spillover from Los Angeles has mainly affected its surrounding areas like Anaheim or San Bernadino, whilst more distant regions seem to enjoy very limited economic benefits.

Shanghai is a highly developed, rich metropolis within China. Due to lack of distance and easy transport access, dramatic economic development of Shanghai has been able to spread into its surrounding regions within Zhejiang and Jiangsu provinces. Meanwhile, in Anhui province, which is geographically distant from Shanghai, the spread effects generated from Shanghai on the economic growth in Anhui tend to be small and weak. As a result, Anhui remains one of the poorest provinces in China. Similarly, in contrast to the central region, the western region, particularly the hill area, has been facing much longer journeys to the eastern and coastal regions in China. Consequently, economic development in the western region has been slower than in the central region. As Phillips and Yeh (1990) point out, in the context of China, the arguments of the growth engines of coastal regions helping the economic growth of the inland regions and the economic prosperity spreading into other interior provinces from the coastal ones have not proven to any desirable extent.

5.3.3 Evidence against Neoclassical Growth Theory

Many scholars have supported regional divergence theories; their studies demonstrate that many nations have been facing the serious challenge of spatial income disparity between the core and peripheral regions; because of absolute domination by the backwash effects in shaping regional development. For instance, Venezuela's wealth and population are highly concentrated in its capital, Caracas and its surrounding areas (Friedmann, 1966). Pióro's case study (1972) suggests the regional inequality in

Tanzania. According to Pióro's estimation, the national wealth is concentrated on the belt between Lake Victoria and Tanga, which occupies only about 10 percent of total land area of Tanzania. However, 40 percent of national GDP, 25 percent of national industrial production, and 70 percent of electricity production of Tanzania have been generated in this belt.

As far as the developed nations are concerned, the interregional income divergence trend rather than convergence is notable. In terms of net disposable income per capita, for example, in Britain, the overall Gini coefficient was up by 30 percent during the period from 1978 to 1991, which was double the equivalent figure in the 1949-1976 period. In addition, relative to the downtrend trend of income inequality before the 1970s, the Gini coefficient has changed to show a steadily increasing trend in Japan over the last two decades (Cornia, Addison and Kiiski, 2004). According to the figures given by Cornia, Addison and Kiiski (2004, p. 29), in Japan, the Gini coefficient had reached 0.44 in 1993 up by 15 percent from the equivalent figure of 0.3 in the 1970s, in terms of income before taxes or transfers. Similar upward Gini coefficient trends have also been found in terms of the income concept after taxes, before transfers and after taxes, after transfers. Singh and Dhumale (2004) argue that the net disposable income per capita Gini coefficient has demonstrated a significantly increasing trend in Sweden. Atkinson's study (2004) suggests that the nations within the OECD face the challenge of unequal distribution of net disposable income due to unemployment and widening wage variation. Moreover, the empirical research conducted by Cornia, Addison and Kiiski (2004) suggests that the divergence tendency of personal income has been confirmed from the study of the East European nations between 1989 and 1995.

In terms of China, the study conducted by Chen and Fleisher (1996) suffers from some problems. In fact, the evidence of regional economic convergence in China is confined to within the coastal and inland regions. However, the tendency toward economic convergence between the coastal and non-coastal regions across China simply did not occur in any desirable degree. As Chen and Fleisher (1996, p. 154) admit, "This implies that convergence is occurring within the coastal and non-coastal groups, but not between them." To my knowledge, one thing that really matters for China is the spatial disparity reduction between the coastal and inland regions rather than narrowing inequality within each group alone. Without proper evidence demonstrating economic disparity reduction between the rich coastal (developed) and poor interior (under-developed) regions, it is actually hard to claim that China has made dramatic progress in spatial disparity reduction and balanced regional development.

In general, I am not so convinced by the argument made by the neoclassical growth model. Regional economic convergence claimed by the neoclassical theory is highly uncertain. The prediction of regional convergence and spatial inequality reduction remains to be confirmed. The firm argument of regional income convergence claimed by Barro and Sala-i-Martin is not so reliable. How long will regional convergence

take based on the two percent convergence rate; one generation, two generations or even longer? The research by Martin and Sunley (1998) has clearly raised this concern. They argue that even in America, which is the most representative case supporting the two percent growth rate of regional convergence, it will take about 35 years in order to reduce just 50 percent of its initial spatial inequality. Moreover, for Europe, it will take 70 years to reduce half of its initial degree of spatial economic inequality. Due to the debatable data on regional inequality reduction within the developed countries, Martin and Sunley (1998) have raised strong doubts on the regional convergence prediction proposed by the neoclassical economists. Moreover, Fingleton (1999, p. 29) further maintains that,

... using a conditional autnormal model, with the effect that there is even less evidence for conditional convergence and for the empirical validity of the neoclassical growth model.

The tendency of regional income divergence between the poor and rich countries has been found by Quah (1992). Based on his alternative theoretical model, Quah (1996) provides more detailed study of regional economic polarization. Instead of the regional economic convergence claimed by the neoclassical growth model, Quah (1996) argues that the developed regions tend to become more prosperous; while in contrast, the under-developed regions tend to become more backward and poorer.

Barro and Sala-i-Martin seem to be over-optimistic about regional convergence, given the lack of sufficient undisputed evidence. For example, in the regions within the European Union, Fingleton (1999) finds that regional convergence reflected in the per capita output remains quite weak. Moreover, the study conducted by Martin and Sunley (1998) suggests regional economic clustering in the northern regions within Europe. The rapid economic development in the northern regions is impressive; by contrast, the under-developed places with slow-growth tend to concentrate in the Mediterranean region of Europe. By referring to the case studies of the developed countries, they disagree with the argument of regional convergence proposed by the neoclassical economists. Furthermore, based on the stochastic model and empirical estimation of 102 countries during the period from 1960 to 1989, Lee, Pesaran and Smith (1997) claim that the regional economic divergence across the world rather than spatial convergence persists due to the higher technology growth in the developed countries (e.g. the OECD).

One of the important limitations of the neoclassical growth model is perhaps due to its explanation of exogenous technological diffusion. The neoclassical model has failed to provide a clear and convincing argument regarding the exogenous technological effect of stimulating long-term economic growth. Aghion and Howitt (1998) point out this shortcoming. Similarly, Fingleton's study (1999) has further suggested that the key causing forces of technical innovation and knowledge spillover have been left unaddressed by the neoclassical economists.

Moreover, once again, as far as East Asia is concerned, the so-called 'East Asian Economic Miracle' has probably posed the biggest challenge to the reliability of the neoclassical growth model. As discussed earlier in this chapter, according to the neoclassical model, the long-run economic growth in a region must be based on technological innovation and technology diffusion. Due to the effect of diminishing returns to capital scale, capital investment cannot stimulate long-run economic development. The development strategy depending on capital alone will inevitably cause regional growth to cease in the future. The only factors accounting for long-term economic development are technological progress and total factor productivity. As Fingleton (1999, p. 7) states, "A key feature of neoclassical theory is that sustained growth of output per capita does not occur unless there are shifts in the production function resulting from exogenously determined technological progress, so the rate of technological progress determines the long-run rate of growth." The case of East Asian countries has been cited by the neoclassical economists to support their argument. According to the arguments made by the neoclassical growth theory, the determining factors contributing to the East Asian economic development are high capital investment and rich labour resource rather than technological progress and knowledge innovation as shown in the total factor productivity. Akhand and Gupta's study (2006) clearly supports this argument.

Therefore, owing to the effect of diminishing return to capital, the economy of East Asia, which depends on domestic savings and investment, will inevitably slow down and cease to grow in the long term. In his article, '*The Myth of Asia's Miracle*' published in *Foreign Affairs*, Krugman (1994) has given his strong and impressive criticisms of the economic miracle in East Asia (particularly Singapore). Krugman firmly argues that the huge inputs of fiscal capital and human resource have entirely accounted for the rapid growth in East Asia; there was almost no contribution made by technological innovation and productivity efficiency to East Asian development during the high-speed economic growth period between 1965 and 1993. He asserts that the economic growth in this region is just a copy of Stalin's Soviet Union. In a comparison between Soviet Union and East Asia, Krugman claims that their economic development stories are almost the same. This is in precise terms, the input-driven type of economic development. According to the argument made by Krugman, Lee Kuan Yew's Singapore is a classic example supporting the input-oriented growth; Krugman (1994, p. 67) refers to it as the "growth achieved purely through mobilization of resources."

Moreover, according to the figures presented earlier in Figs 5.2 and 5.3, the rapid economic growth in many East Asian countries seems to largely rely on huge capital investment and rich labour force, in areas such as Hong Kong (China), South Korea and Singapore. Young (1995) suggests that the growth rate of total factor productivity (TFP) has played an insignificant role in the high-speed growth of East Asia. For example, relative to the OECD and some other developing states, the TFP growth rates were only 2.3 and 0.2 percent in Hong Kong (China) and Singapore respectively. Young (1995, p. 641) claims that,

While the growth of output and manufacturing exports in the newly industrializing countries of East Asia is virtually unprecedented, the growth of total factor productivity in these economies is not.

In addition, Kim and Lau (1994, cited in Akhand and Gupta, 2006) argue that the inputs of capital and labour have entirely accounted for the rapid economic growth in East Asia during the post-war period between 1966 and 1990. Capital and labour are the key factors stimulating the East Asian development. The contribution generated from the technological factor can be virtually ignored. By contrast, the relative contribution made by technological progress to economic growth in many western countries was much more significant than in East Asia during the similar period. For example, technical progress contributed to 65 percent of the annual GDP growth in France between 1957 and 1990 respectively (see Table 5.1 below). Similarly, Aghion and Howitt (1998, p. 404) conclude, “That much of East Asian growth can be attributed to factor accumulation, in turn, seems to imply that technological progress is relatively unimportant as a source of growth.”

Table 5.1: Sources of Economic Growth in an Estimation Model without Addressing Human Capital in Selected Regions and Countries 1948-1990

Country	Time period	GDP ^a	Relative contribution ^b of:		
			Capital	Labor	Technical progress
East Asian NICs:					
Hong Kong	1966-90	7.8	74	26	0
Korea, Republic of	1960-90	8.6	80	20	0
Singapore	1964-90	8.9	68	32	0
Taiwan, China	1953-90	8.7	85	15	0
G-5 countries:					
France	1957-90	3.7	37	-1	65
Japan	1957-90	6.7	56	5	39
Germany	1960-90	3.2	40	-7	66
United Kingdom	1957-90	2.5	39	4	57
United States	1948-90	3.1	26	26	49

Source: Adapted from tables 3.1 and 7.2, Kim and Lau (1994).

Note

a Average annual growth rate of real GDP (percentage points).

b Relative contributions of inputs as percentage of output growth.

Reproduced from Akhand and Gupta (2006, pp. 44)

Similarly, when comparing the high-speed growth in East Asia and the golden development age of Europe, except for the United Kingdom, Crafts (1999) points out that the TFP is the key source explaining the fast economic growth in Europe during the period 1950-73. However, as far as East Asia is concerned, Crafts suggests that capital and labour accumulation had contributed to its rapid economic development, but not to the TFP. For example, in Italy and West Germany, the relative annual

growth percentages of TFP were up to 3.2 and 3.3 between 1950 and 1973; while the equivalent figures were only 1.5, 0.9 and 0.8 in South Korea, Malaysia and Indonesia between 1960 and 1994 respectively (1999, p. 150).

However, the arguments made by the neoclassical economists have been challenged by numerous studies, particularly in recent research. For example, the empirical estimation made by Klenow and Rodriguez-Clare (2001) illustrates that the important contribution made by productivity growth in stimulating the output per worker in East Asia should not be undervalued. In general, the growth of TFP was extremely impressive in East Asia. For example, the growth rates of adjusted total factor productivity were 3.7, 2.5 and 3.5 percent in Hong Kong (China), South Korea and Taiwan (China) respectively during the period of high-speed economic development. They argue that the TFP gains accounted for the substantial growth of output per worker in these economies. Moreover, relative to America, Sarel's empirical research (1997) suggests that the annual growth rates of TFP were impressive in many East Asian countries during the period between 1978 and 1996. For example, relative to the equivalent figure of 0.3 percent in America, the average annual growth of TFP were up to 2.0, 2.3 and 2.0 percent in Thailand, Singapore and Malaysia. (Sarel, 1997, p. 32)

As far as China is concerned, one of the recent studies made by Bosworth and Collins (2008), highlight that this country has experienced dramatic TFP growth during the reform period. In terms of the industrial output per worker, the impressive growth rate has been largely due to the growth of TFP and capital per worker since 1993. According to the data presented in Table 5.2 below, the contribution made by the TFP has accounted for almost 50 percent of the average annual growth of output per worker since 1978. By contrast, the growth of TFP in the other East Asian countries was significantly lower than the equivalent figures in China during the similar period. In terms of the period between 1980 and 2003, except for China, the contribution generated from the TFP growth to the output per worker in East Asia was merely 0.9 percent. In general, Bosworth and Collins (2008, p. 53) state that,

This feature sets both China and India apart from the East Asian miracle of the 1970s and 1980s, which was more heavily based on investments in physical capital. In addition, China stands out for the sheer magnitude of its gains in total factor productivity.

Table 5.2: Sources of Economic Growth in China, India, and East Asia during the period 1978-2004 (Annual percentage rate of change)

Period/country		Output	Employment	Output per worker	Contribution to output per worker of		
					Physical capital	Education	Total factor productivity
1978-2004	China	9.3	2.0	7.3	3.2	0.3	3.6
	India	5.4	2.0	3.3	1.3	0.4	1.6

1978-1993	China	8.9	2.5	6.4	2.4	0.4	3.5
	India	4.5	2.1	2.4	0.9	0.3	1.1
1993-2004	China	9.7	1.2	8.5	4.2	0.3	3.9
	India	6.5	1.9	4.6	1.8	0.4	2.3
East Asia excluding China							
1960-1980		7.0	3.0	4.0	2.2	0.5	1.2
1980-2003		6.1	2.4	3.7	2.2	0.5	0.9
1980-1993		7.3	2.7	4.6	2.6	0.6	1.4
1993-2003		4.5	2.0	2.5	1.8	0.5	0.3

Source: Authors' estimates as described in text; Bosworth and Collins (2003).

Notes:

1. The employment series is a census-comparable concept for both China and India. The East Asia comparison includes Indonesia, South Korea, Malaysia, Philippines, Singapore, Taiwan, and Thailand.
2. Growth rates may not sum due to rounding.

Reproduced from Bosworth and Collins (2008, pp. 49)

Previous empirical research seems to produce very contradictory findings regarding the TFP growth and contribution made by technological progress to East Asia. This is properly due to differences of specific research methodology and assumption, statistics data and the various time-series studies (Ákhand and Gupta, 2006). They suggest that,

It appears that by reworking the data one can arrive at almost any plausible conclusion. This highlights the general fragility of inferences drawn in the productivity growth empirics literature devoted to explaining the East Asian Miracle. (2006, p. 54)

From my point of view, due to dispute over measurement approaches and component weighting, the exact contribution of TFP to the East Asian economic development is still arguable. Capital and labour force are the two important factors explaining the rapid economic growth in this region for most of the past three decades. However, it is not justified to reach the oversimplified conclusion of the neoclassical economists, who allege that the contribution generated from technological progress to the economic growth in East Asia is insignificant or even can be entirely ignored. The precise contribution made by technological innovation and productivity to the East Asian economic success remains unclear and still needs far more empirical research. Nevertheless, it is actually hard to accept the neoclassical view that technological progress has not made any contribution to local economic growth and industrial productivity enhancement in East Asia. Apparently, this is not the case. In fact, many East Asian countries have made impressive achievements in terms of technological progress. Owing to the diminishing return to capital, the neoclassical economists predicted that the economic growth in East Asia which depends on high capital investment would eventually slowdown and even entirely stagnate. However, in

contrast to many other regions in the world, even though the current average growth rate is lower than the equivalent figure in the past high-speed development period, the economic development in East Asia is still outstanding; this is especially true of China. It is reasonable to argue that productivity growth is important for stimulating the economic growth in this region. Technological progress has been important for East Asian economic success. In terms of this aspect, the views suggested by the neoclassical theory have certain shortcomings. According to the neoclassical growth model, the poor nations with the high return of capital investment would eventually catch up with the rich in terms of income level under the condition of equal growth of labor, technological progress and domestic savings (Pack, 1994). Pack further claims that the regional income convergence suggested by the neoclassical model might not happen due to its invalid main argument of decreasing return to capital, which is caused by the economic externalities and productivity growth (supposing knowledge and technological innovation can be regarded as one kind of capital). Significantly, even if they recognize that factor accumulation is crucial, Akhand and Gupta (2006) implicitly hint at the important role played by the TFP in understanding the economic success in East Asia.

In my view, Barro and Sala-i-Martin's studies rely too much on developed countries like America. Relative to the developed countries, the regional development situation within many under-developed countries is quite different. Consequently, the conditions for achieving regional convergence might be dramatically different from in the developed countries studied by Barro and Sala-i-Martin. For instance, in Latin America, Bengoa and Sanchez-Robles's research (2003) has raised the serious concern about the validity of the neoclassical growth model. In contrast to the arguments made by the neoclassical growth theory regarding diminishing capital returns and the insignificant role of capital investment in the long-run development, they point out that foreign capital investment is always closely linked to economic growth. The FDI played a positive role in regional economic development in Latin America. They suggest that capital investment is still an essential and key factor in boosting economic growth.

Moreover, suppose the conditions of sufficient human capital and technology diffusion are guaranteed, will the tendency of regional convergence naturally occur? I strongly doubt it. Besides the human and technological factors, are there any other conditions required essential to achievement of regional convergence? In addition, the geographical effects on spatial economic development and regional divergence seem to be completely neglected by Barro and Sala-i-Martin's neoclassical growth model. In fact, even Sala-i-Martin himself recognizes the problems with regional convergence and the limitations of his study. He admits that regional economic divergence rather than convergence persists in many countries across the world. As Sala-i-Martin (1996, p. 1034) admits,

First, the cross-country distribution of world GDP between 1960 and 1990 did not shrink, and poor countries have not grown faster than rich ones.

Using the classical terminology, in our world there is no σ -convergence and there is no absolute β -convergence.

Therefore, it is reasonable to claim that the arguing power of neoclassical growth model is extremely limited, and the prediction made by this model is problematic and ambiguous. Conditional convergence of regional income is not supported by world evidence. In fact, numerous case studies reveal that regional income divergence is the highly possible development tendency of spatial economic growth.

5.4 Conclusion

In this chapter, the literature on the main theories of regional income convergence has been assessed. The world evidence regarding the developing trend of spatial inequality has also been discussed. In general, Kuznets's research unveils the U-shaped pattern of regional development process. During the early development stage, the process of regional economic growth would firstly show the trends of economic agglomeration and widening spatial income disparity, followed by spatial economic convergence and balanced regional development in the high development period. Some parts of Kuznets's inverted U-curve inequality model are justified. It suggests that geographical influence on regional divergence should not be underestimated. Geography makes a difference to regional economic development. In addition, Kuznets argues that the journey toward regional convergence will take a very long time. However, Kuznets' hypothesis suffers from serious limitations and I am not convinced by his hypothesis. Kuznets's answers to certain important issues are either inappropriate or have not been fully researched. Kuznets's theory has been facing strong criticism. More significantly, the majority of the world evidence does not support his classic hypothesis of the inverted U-curve pattern of regional disparity.

The conditional convergence claimed by the neoclassical growth model suggests that spatial inequality reduction and regional convergence will not necessarily happen. If the under-developed regions do not have a rich labour force, the economic divergence between the developed and under-developed regions might become even worse. In addition, the neoclassical scholars argue that technological progress and knowledge spillover is the key to achieving sustainable economic growth. However, without question, the neoclassical growth model has certain limitations, and its key argument is not justified. The prediction of regional convergence and spatial inequality reduction made by the neoclassical scholars is not convincing and remains to be confirmed. Barro and Sala-i-Martin are over-optimistic about the assumption of regional income convergence in the long term. A lack of sufficient and undisputed evidence makes the neoclassical growth model theoretically illogical and empirically weak. Even if the neoclassical view is correct, based on the two percent growth rate of regional convergence, the tendency toward regional income convergence will take many years. Moreover, the neoclassical model has been unable to provide convincing

argument with regard to the exogenous technological factor in stimulating regional economic growth. What is the actual functioning mechanism for the technology spillover in boosting regional development process? These questions have not been properly addressed by the neoclassical growth theory.

World evidence has largely rejected the argument of the neoclassical growth scholars regarding the development tendency of regional income convergence. In particular, the high-speed economic development in the East Asia has posed the biggest challenge to neoclassical economists. Their claim that economic growth in East Asia inevitably will slow down and even halt in the long term, has not been supported. In contrast to other regions in the world, the economic growth rate in East Asia is still impressive, for example, in China. According to arguments proposed by the neoclassical model, instead of technological progress and the strong state, the rapid economic growth in this region has been entirely based on the accumulation of physical capital and labor force. However, the postwar development experience in the East Asian countries seems not to support this argument. On the one hand, the strong state accounts for the economic success of East Asia; particularly with regard to state-designated industrial policies. On the other hand, technological progress as shown in the growth of TFP has made a difference to economic development in this region. Insistence on regional economic growth's dependence on capital investment and physical labor alone is simply not supportable.

What about the Chinese evidence? Which changing trend does China represent: convergent or divergent? As far as Guangdong is concerned, does the development tendency of regional disparity follow the inverted U-curve hypothesis suggested by Kuznets and conditional convergence proposed by the newclassical growth scholars? Has regional inequality been showing a convergent trend during the reform period? Does inequality trend evidence in this southern province differ from the majority of the world? What is the unique perspective in terms of the case of Guangdong? Chapter Nine offers detailed discussion of this issue. Nevertheless, before specifically discussing Guangdong, it is essential to address the broad evidence in China regarding the inequality of development tendency and also the influence of geography on economic growth.

Chapter 6 Geography, Economic Growth and Regional Inequality in China

6.1 Introduction

The central theme of my research is to investigate the changing tendency of regional inequalities within Guangdong Province and the role played by geography in influencing spatial disparity. However, before addressing Guangdong, it is necessary to discuss the evidence of other Chinese regions with regard to the impact of geography on regional economic development, and the development tendency of regional inequality. This research will then be in a better position to make an in-depth study of the regional development phenomenon in Guangdong. This chapter attempts to present a general literature survey on the effects of geography on Chinese economic growth, and the changing trend of spatial income inequality in this nation. In precise terms, what is the impact of geographical location and market proximity on industrial and economic growth in China? What is the role played by industrial cluster and agglomerative economies in affecting regional development? What is the developing trend of spatial income inequality in China? These key questions are addressed by this study.

The structure of this chapter is as follows. Section Two presents the contribution made by traditional geographical theories to assessment of regional economic development in China. The first part of this section emphasizes the advantage of good location; the second part discusses the effect of transportation costs. Section Three focuses on the analysis of industrial cluster and economic agglomeration. The importance of historical factors in developing the local industrial cluster is also analyzed in this section. Section Four offers analysis of the tendency of regional disparity in income. The developing trend of spatial inequality in China is the key issue addressed in this section. Numerous case studies will be cited to test the argument of regional income convergence and inverted U-shaped model of regional income disparity made by Kuznets and the neoclassical economists. Section Five discusses the macroregion model developed by Skinner. Finally, the chapter draws conclusion in Section Six.

6.2 China and Traditional Economic Geography

6.2.1 Geographical Advantage and Coastal Proximity

Sung, et al, (1995) claim that the coastal regions themselves and the regions in proximity to the coast or ocean-navigable rivers in China have achieved dramatic economic development. Wang and Hu (1999) summarize the importance of

geographical location and distance in influencing regional prosperity. They offer a detailed discussion on the causes of uneven regional development in China. In particular, Wang and Hu highlight that one of the key factors contributing to the spatial economic disparity between the rich and poor regions is geographical location.

The Chinese city of Shanghai, located within the Yangtze River Delta Region, is the largest economic centre and industrial powerhouse of China. However, it is important to note that Shanghai is also a key coastal metropolis located near to both the Pacific Ocean and ocean-navigable Yangtze River. The geographical advantages of coastal proximity and good harbour are the main factors causing low transportation costs, and contributing to the rapid industrial and economic development in this region. The advantage of transportation hub further reduces the transportation costs in Shanghai and reinforces its economic leading position. Zhao (1994, p. 224) maintains that, "... Shanghai is actually served by a network of waterways, roadways, air routes, and railways." Moreover, Sun stresses the striking correlation between the inherited good harbour in Shanghai and its economic power in China. Sun (1988, p. 346) points out that,

As China's largest hub of river, lake, and land transport, with large areas of available land, Shanghai is unmatched by other harbors in terms of its favorable conditions.

Beijing, Tianjin and Dalian, are three developed and prosperous regions, all located within the Bohai Gulf Region; the ocean-navigable Bohai Bay provides these three municipalities with good and fast access to the coast (Sun, 1988; Zhou, et al, 2002). The empirical research of Démurger (2001) also shows that geographical location plays an important role in the regional economic performance of China by using panel data from 24 provinces. The Yangtze River Delta Region; one of the most dynamic and prosperous regions in China; its development is also strongly associated with its proximity to the ocean-navigable Yangtze river.

However, not every region in China has inherited the geographical advantages of the eastern region. In fact, a large portion of China's land is located in the central and western regions which are distant from both the sea and ocean-navigable rivers. These regions suffer from unfavourable geographical features; many interior regions have achieved relatively slow economic growth and little improvement of people's living standards so far. This is probably because these regions are remote, landlocked, or have access blocked by mountains. Wang and Hu's study (1999) offers a detailed discussion regarding the extremely mountainous geography in the majority of regions within China (up to two thirds of Chinese land).

In fact, the majority of the remote and landlocked regions are closely associated with extreme geography and high transportation costs. Therefore, it is reasonable to claim that one of the crucial factors causing the slow economic growth and low personal income in these under-developed Chinese regions is high transportation costs mainly

affected by geographical location. For example, in Lhasa, Xizang Autonomous Region (Tibet), mountains, extreme remoteness and long distance to the core eastern region in China, inevitably lead to high transportation costs. Tregear (1970) suggests that the motor road to Lhasa is tough and inaccessible; vehicles need to pass through numerous deep valleys and tall snowy mountains. Moreover, Pannell and Ma (1983) highlight the highland and mountainous features of topography in Xizang, which causes low population density and lack of investment incentives at both the domestic and international levels. Consequently, it leads to slow economic growth. Qinghai – Xizang Plateau is, in fact, the highest tableland³¹ in the world, and is surrounded by rocky mountain ranges. The study conducted by Grunfeld (1996, p. 7) highlights the extremely mountainous topography,

... Tibet is a stunning place. Predominantly a plateau averaging 3,600 m. (12,000 ft.) above sea-level, its landscape includes not only snow-covered mountains but also glaciers and green forest, grasslands and salt lakes. It is surrounded by mountains ...

Moreover, the geographical disadvantages in Xizang are reflected in its extreme climate. This place is a classic example demonstrating the severe negative effects of extreme climate on agricultural, industrial and economic development. Aridity and air rarity make this a hostile environment for mass agricultural production or development of modern industry and foreign capital; therefore, it may well suffer from slow economic growth and regional backwardness. Numerous previous studies have emphasized the extreme climate in Xizang. An early study by a Russian scholar points out the poor climate in this remote region. In particular, Zaychikov (1969) highlights the air rarity in Xizang. Similarly, Tregear (1970) also discusses the extreme climate in Xizang. In particular, Tregear stresses that natural phenomena such as poor climate, tall mountains and desert, all make the accessibility and communication of Xizang with other regions extremely difficult and dangerous. In general, the double impacts of extreme climate and mountainous characteristic of topography lead to perpetuation of slow growth and economic backwardness in Xizang. More significantly, the poor climate indicates that economic prosperity might never be achieved in this place. In addition, Sivin et al, (1988) discuss the extremely cold and dry weather in Qinghai province. They maintain that poor climatic factors like long winters with low temperature, arid land, and insufficient rainfall (only 25 mm a year) are the important reasons for the economic underdevelopment in Qinghai.

In my opinion, as far as Xizang is concerned, physical geography has affected economic growth. In precise terms, the inherited locational disadvantage and extreme

³¹ “The Chinghai-Tibet Plateau in southwest China embraces mainly the Tibet Autonomous Region, Chinghai Province and western Szechuan Province. It is the highest tableland in the world, and has an area of 2.2 million square kilometers and an average elevation of over 4,000 meters above sea level. The Chinghai-Tibet Plateau, with its series of glaciers and immense snow-capped peaks, is known as ‘Roof of the World.’” (Chin, 1978, p. 44)

climate account for economic underdevelopment. Industrial prosperity and rapid economic development are hard to realize in this region. The spatial economic gap between the developed Chinese regions like Shanghai and under-developed Xizang will continue to widen for many years to come.

A region such as the south-western province of Sichuan, which is relatively backward in comparison to the rich eastern provinces, might be largely affected by inherited poor geography. Tall, rocky mountains negatively affect the communication between Sichuan and the rest of China; and this disadvantage reduces opportunities for economic cooperation between Sichuan and the other regions, discourages foreign investment and prevents active participation in international trade. Apart from the sea route along the Yangtze River, Sivin, et al, (1988) highlight that the mountains located in Sichuan hinder its access to the other regions. In addition, their research also suggests that the extremely mountainous and remote nature of its topography has generated negative effects on economic growth in Guizhou; and makes its economic backwardness difficult to overcome. Sivin, et al, demonstrate the linkage between interior landscape and economic backwardness in Guizhou.

In China, relative to the inland regions, the coastal regions are more favourable for economic development. The correlation between coastal location and rapid economic growth is close. In contrast to the interior central and western regions, the eastern region located near to the sea or navigable rivers has achieved much faster trade and industrial development, and higher rate of economic growth.

6.2.2 Transportation Costs and Economic Development

In the context of pre-1949 China, the role of backward transportation and other poor infrastructure in preventing the industrial and economic development is evident. Buchanan (1966) argues the negative impact of transportation shortage on national economic and political integration of China. Particularly, he stresses that insufficient transportation facilities in the interior central and western regions are even prominent; many places located in these regions have failed to develop any kind of modern transportation system. Buchanan (1966, p. 63) states that,

Such conditions [**insufficient transportation system**] made the development of a national economy impossible; the country remained broken up into a series of more or less developed areas, separated from one another by a stagnant sea of subsistence production ... (Bold content added)

Today, many of the least developed regions are still suffering from a lack of efficient transportation facilities, which accounts for their slow industrial and economic growth. The study of Chinese geography illustrates that many mountainous and under-developed regions have rich reserves of natural resources. Natural resources such as coal (e.g. inland Shanxi Province) and oil (inland Xinjiang and Heilongjiang)

are located in many parts of China. Li's study (1990) presents a detailed discussion on this issue. Owing to the inherited mountainous and inland topography with long distance to large markets, and lack of efficient transportation facilities, the least developed regions have had difficulty realizing the economic potential of valuable natural resources. Due to high transportation costs and geographical problems, the foreign investors have lacked sufficient incentives to invest in these regions by exploiting these natural resource reserves. Li (1990) points out the oil case of Xinjiang. Owing to the poor weather, extreme geography (e.g. Gobi Desert) and long distance to the coast, the exploitation and transportation of oil is too expensive.

In Xinjiang, a remote and resource-rich region located in western China, the long distance to the eastern region probably is one of the important factors causing its slow economic growth. Sivin, et al, (1988) describe the mountainous landscape in this place. In addition, the poor and backward interregional transportation facilities further impede local economic development. For the commercial enterprises, the transport of raw industrial materials and agricultural produce like cotton and tomato from Xinjiang to the main Chinese and overseas markets is costly; moreover, high transportation costs discourage firms from setting up manufacturing industries in this place. Démurger (2001, p. 115) argues that,

An explanation [of **poor economic performance**] can be found in its underdevelopment in terms of transport facilities which, in a strongly agricultural province, makes goods transportation to and from markets extremely difficult and expensive. (Bold content added)

In addition, Wiemer (2004) states the difficulty in attracting the foreign capital to Xinjiang due to distance from both large consumer markets and the countries with rich investment capital. Li's study (1990) suggests that not only products manufactured in Xinjiang, but also its home-grown fruits and vegetables, are at a price disadvantage in regional competition. Although the various fruits grown in Xinjiang enjoy the reputation of good quality and delicious flavour: like Hami melon, it is hard to compete with similar fruits grown in other regions due mainly to the inherited geographical remoteness and long distance to the core market. The study made by Wang and Bai (1991) clearly suggests this point.

From my point of view, although its geographical disadvantages have been reduced dramatically by the heavy state infrastructure investment during the reform period, without question, the local industries and economy in Xinjiang still suffer from the effects of remote topography and long distance to the core market. The inherited extreme geography makes it difficult to develop export-oriented manufacturing industries and attract foreign capital. According to some principles of market economy – low production costs and consumer market proximity – when faced with low profit return and high investment risk, the potential investors have lacked the desire to invest in Xinjiang. This is why Xinjiang remains relatively one of the economically poor and slow-growing regions in China. The political factors (e.g.

social unrest and the conflict between the Han and other ethnic groups) are of course other factors explaining the economic underdevelopment in Xinjiang; however, without any doubt, geography is the main cause. The opening up of Xinjiang has been largely blocked by its own geographical remoteness. In terms of this point, Gifford (2007) seems to greatly underestimate the effect of geographical isolation on local economic growth in Xinjiang. He inappropriately alleges that distance would not be an issue for the economic development of Xinjiang.

In general, due to their strong economic growth and increasing local tax base, the coastal and eastern regions have achieved rapid development of transportation. The local government can put more public investment into infrastructure construction. For example, the development of modern transportation facilities of railway, road, water and air has significantly contributed to the industrial and agricultural development in Dalian (Sivin, et al, 1988). In addition, the study of Chinese transportation development conducted by Wang and Hu (1999) further suggests that the advanced and modern transportation facilities have provided huge benefits for the coastal and eastern regions. These facilities enables the various raw materials, products and passengers to be transited with great efficiency; meanwhile, the western and central regions have been facing the double challenges of both shortage and backwardness of local transportation facilities. Relative to the eastern region, Wang and Hu argue that the central and western regions have much less developed transportation facilities and lower density of transportation network. This is why Démurger (2001) firmly argues that infrastructure improvement in the rural and remote areas within China is so important for stimulating economic development.

Modern transportation system and efficient transportation facilities are the crucial factors shaping regional economic development and social progress. As Skinner (1965a, p. 219) points out,

... the fate of particular markets on a modernizing landscape is essentially dependent on the spatial patterning and temporal sequence of transport modernization

In his series of articles during the 1960s, William Skinner provided detailed research into the history of market formation and social structure in rural China. Significantly, Skinner (1964; 1965a; 1965b) presents a model which stresses the importance of transportation construction to market development and formation of city and society. According to the regional model proposed by Skinner, the formation of the large markets and commercial towns has largely depended on transportation infrastructure. The conditions of local accessibility and goods mobility are crucial for market and commercial development. Skinner (1964) suggests that geographical distance can limit the potential demand for the private agents and dependent regions of the core market. Skinner (1964) indicates that the market periodicity in the under-developed peripheral regions relies on the status and functioning condition of transport. As he states,

Thus the periodicity of markets in traditional agrarian societies is, in the last analysis, a function of the relatively primitive state of transport. (1964, p. 11)

In the context of China, Skinner (1965a) argues that the developed regions and commercial centres have better transportation facilities, and better infrastructure than the under-developed agricultural periphery. Therefore these areas tend to achieve faster economic growth and social development. Skinner (1965a) found that modern transportation facilities had concentrated in the urban and core regions, for example, railway construction (Nanjing – Shanghai and Qingdao – Jinan), public highways (Changshou – Shanghai and Beijing – Tianjin) and the introduction of steamships (Hankou – Nanjing and Xiamen – Shantou). Skinner (1965a) seems to firmly believe that the improved transportation system in the urban centres has substantially reduced the transportation costs between the various urban and core regions within China. In addition, although he recognizes the improvement of transportation in relatively under-developed places (e.g. highway construction in all counties within Sichuan and Guizhou provinces), Skinner's regional model (1965b) still draws the following conclusion. It is probably because of insufficient and backward infrastructure connecting the developed places with under-developed areas that the economically poor periphery is unable to change its industrial structure, develop modern industries and speed up local economic growth. Therefore, the under-developed periphery has largely remained as a feature of traditional agricultural society. A later study conducted by Skinner (1985, p. 412) further suggests that,

The recognition is growing in China that the infrastructure of rural marketing is grossly inadequate and must be systematically upgraded and modernized.

As far as modern transportation facilities are concerned, significantly, railroad construction has brought substantial economic benefits to regional development. Liang's empirical study (1982) shows that the railroad has stimulated agricultural growth in China. Chin (1978) gives some detailed discussion regarding railroad construction development in China and its significance in stimulating interregional commercial interaction and regional economic growth. For example, the regions along the line of the Lanzhou – Xinjiang railway (*Lanxin line*) have enjoyed the benefits contributed by this railway effect. Before the construction of this railway, many regions, such as Yumen and Urumqi, were very difficult to access and suffered from geographical remoteness. Moreover, the importance of Baotou – Lanzhou railway in stimulating the economic development in Inner Mongolia should not be underestimated. The construction of this railway has made the rich metal resource in Baotou become exploitable, and transformed this place into a crucial iron and steel production centre in China. The traditionally underdeveloped Hetao Area located in Inner Mongolia also shares the economic benefits brought by the Baotou – Lanzhou railway. By contrast, lack of rail network hindered the commercial communication of

the southwest provinces (Yunnan, Guizhou and Sichuan) with the outside world before 1949. However, railway development in these regions has increased substantially since 1949 due to the construction of Chengdu – Chongqing, and Baoji – Chengdu, key railways (Chin, 1978). Moreover, Chin highlights the importance of Yingtan – Xiamen railway in overcoming the geographical disadvantage of remote and mountainous topography in Fujian Province.

Similarly, Sun's study (1988) suggests that the construction of Yingtan – Xiamen railway linked inaccessible regions with the rest of Fujian. Rich forest resources in this province became exploitable. In addition, new industrial centres like Zhangzhou and Sanming have flourished in Fujian mainly due to construction of this railway. Sun (1988) discusses the role played by transportation in boosting Chinese economic growth; in particular, he highlights the effect of railroad construction. Sun maintains that the impact of railroad on economic development is remarkable. It contributes to the transportation of raw agricultural and industrial materials to the place where they are required, more quickly and at lower cost. Sun points out two specific railroads of Yingtan – Xiamen and Lianyungang – Lanzhou (*Longhai line*). Yingtan – Xiamen rail line is crucial to Fujian's economic development; while the economic growth in the western region has significantly benefited from the *Longhai line*. By contrast, Wang and Bai (1991) claim the transportation problem is restricting regional commercial communication and local economic growth in Xizang, as they state,

Bearing this ratio in mind, if we turn to Tibet which has no railways and so relies totally on the road system, and add on the costs of high altitude oxygen deficiency, fuel and spare parts, the restrictions placed on commodity exchange by transportation difficulties are even more evident. (1991, p. 138)

Since 2006, with the successful construction of highland railway, Xizang has been connected with the other Chinese regions by rail for the first time. Although there are some criticisms regarding local environmental pollution and Han migration brought by this railway, the economic development potential of Xizang contributed by the construction of this railway should not be ignored. In fact, this railway has weakened the geographical obstacle of remoteness in this most mountainous region of China. In addition, the railway can substantially reduce the high costs of interregional transportation and strengthen the regional commercial linkages between Xizang and the outside world.

From the historical angle, Huenemann (1984) conducted an important research on railroad development in China before 1949. Huenemann's research suggests that railway construction has brought some benefits for Chinese economic development; this is especially true of the interior regions without access to water transport. Moreover, he points out that the positive effects of the railroad on passenger traffic should not be overlooked. In terms of the contribution made by the railroad to national income growth, Huenemann claims that around an extra 193 million *yuan*

was generated from the operation of railroads in China. Moreover, it was the foreign companies which invested in the Chinese railroad; the Chinese government did not need to finance these railroad construction projects. In general, by contributing to transportation productivity, Huenemann (1984) indicates that the economic benefits brought by the railway are enormous.

However, significantly, the geographical advantage of coastal proximity is perhaps not the only key factor contributing to economic and industrial development in the developed regions. In fact, the advantage of coastal location alone is not sufficient to stimulate economic growth. For example, Wuhan, which is an important national transportation hub located in the central region, has inherited the favored advantages of geographical location and low transportation costs. Wuhan is located on the navigable Yangtze River, which provides this region with good accessibility. In addition, numerous important national railway lines for freight and passenger transportation, like the Beijing-Guangzhou (*Jingguang*) line, meet at Wuhan. Wuhan is also an important inland seaport. Li (1990, p. 80) argues that, "... Wuhan is well known as a 'thoroughfare to nine provinces' and has long been an important commercial centre." Moreover, Sun's study (1988) illustrates the transportation advantage of Wuhan.

Nevertheless, the advantages of location and transportation hub have not contributed to rapid economic development in Wuhan. Its growth is not only slower than many other eastern municipalities, such as Qingdao and Tianjin, but also some central Chinese municipalities like Zhengzhou and Changsha. In fact, the economic and social growth in Wuhan remains generally lower than many other Chinese regions. In terms of Démurger's study (2001), she found that even if some Chinese provinces inherited geographical advantages and good transport access to core regions along the Yangtze River, unlike the eastern provinces, they have not achieved rapid economic growth or gained prosperity. Démurger (2001, p. 114) points out that,

The neighboring provinces along the Yangtze River are hardly industrialized despite their natural resources endowments, in terms of both coal and hydraulic resources, and a relatively favorable geographic position.

The evidence of some Chinese regions indicates that locational advantage and low transportation costs might not be the only key pre-conditions for achieving rapid economic growth. Economic development is also affected by the spillover effect of agglomeration and other external forces. Besides ideal location and low transportation costs, economic growth also goes hand in hand with industrial clusters and economic scale. The advantage of good accessibility and low transportation cost would not necessarily lead to economic growth in an under-developed region. Economic agglomeration is an important influence on geographical distribution of industries and economic activities. As far as the agglomeration effect is concerned, Wuhan is further away from Shanghai than Nanjing and Hangzhou by road (the road distances from

Shanghai to Wuhan, Nanjing and Hangzhou are 919, 349 and 213 kilometers respectively). Therefore, Wuhan has enjoyed much less economic benefit from the agglomeration effect spread from Shanghai and its surrounding Yangtze River Delta region.

My general argument is this: it is reasonable to claim that favourable location and low transportation costs are not the only pre-conditions for economic development. Economic development is closely associated with many other factors, such as state, industrial cluster, education and human capital. This is perhaps why some Chinese regions with geographical advantages have failed to achieve industrial prosperity and rapid economic growth. Geographical proximity to the sea, large market and low transportation cost is not the whole story of regional economic growth. However, in general, the Chinese evidence illustrates the fact that location and transportation costs are essential to regional economic growth and spatial distribution of industry and other economic activities. The crucial role played by the industrial cluster and economic agglomeration in boosting the local economic development in China is discussed in the following section.

6.3 Industrial Cluster and Economic Agglomeration in China

Local economic growth and dramatic industrial expansion of the Yangtze River Delta is a good empirical example of agglomeration, as developed by the NEG. This large region, including Shanghai, Zhejiang, and Jiangsu, is one of the most prosperous and developed areas in China. About 10 percent of total GDP and 10 percent of the industrial output value of China are concentrated in the Yangtze River Delta: this is particularly true of Shanghai, which is the largest industrial, trade and financial centre in China. (Chan, cited in Cheng, 1998)

The close economic and trade cooperation between the places located within the Yangtze River Delta has further reinforced the development process of regional economic agglomeration. Tian (1996) claims that the manufacturing products made in Jiangsu and Zhejiang provinces have accounted for the majority of non-locally produced exporting goods in Shanghai. According to the data published by the Centre for Regional Economic Research, Shanghai University of Finance & Economics (2003), around 50 percent of the industrial products made in Shanghai have been transported and sold in Jiangsu and Zhejiang. Equally, Shanghai has also become the key consumer market for goods produced in Jiangsu and Zhejiang. In terms of industrial products, 30 percent and 20 percent of the goods made in Jiangsu and Zhejiang provinces respectively have been consumed in Shanghai. In addition, the domestic direct investment interaction between Shanghai, Jiangsu and Zhejiang is also strong. Shi's study (2004) specifically illustrates the regional economic and trade cooperation between Shanghai and the city-level regions within these two provinces. For example, more than 10 percent of the agricultural goods produced in Jiaying have been purchased by consumers living in Shanghai.

Significantly, in terms of close regional economic cooperation and dramatic development of industrial cluster within the Yangtze River Delta, the central role played by Shanghai is evident. Shanghai is the economic centre of the Yangtze River Delta. The contribution made by Shanghai to local economic growth and emergence of industrial cluster in many regions within the Yangtze Delta is crucial. In particular, the spillover effect, spreading from Shanghai to the surrounding regions, needs to be stressed. Zhang, Xiong and Shen (2002) point out that the regions located in the Lower Yangtze River Delta have benefited significantly from technology, capital, information and industries, all of which spread from Shanghai.

In that sense, without the spillover effects generated from Shanghai, the prosperity of local economy and emergence of regional industrial cluster within the Yangtze River Delta is hard to image. To explain further in the language of new economic geography: owing to geographical proximity, the spillover effects generated from the agglomerative economy in Shanghai play an important role in shaping the regional industrial cluster within the Yangtze Delta.

Numerous Chinese regions demonstrate the emergence and rapid development of industrial cluster. For example, Chongqing is the largest motorcycle production centre in China. This region represents nearly 40 percent of national motorcycle manufacturing (Sonobe, Hu and Otsuka, 2006). Zhejiang is probably the best example to illustrate the strong tendency of regional industrial cluster. Provincial economic development in Zhejiang has benefited substantially from the agglomerative economy. Some scholars have done important work on this eastern province. Guo, et al, (2002) provide rich case studies showing local industrial cluster and economic agglomeration. Yueqing is the key production base of the low-voltage electrical parts industry, accounting for one-third of Chinese manufacture of this product; and its output value accounts for over 70 percent of total industrial output value in Yueqing. Wenzhou is an important centre for the cigarette lighter industry and leather shoe production. The annual shoe production capability of this region is about 600 million pairs, which accounts for 20 percent of total shoe production in China. As far as the metal lighter industry is concerned, there are more than 500 lighter production enterprises located in Wenzhou; the total number produced annually is about 850 million. The lighters made in Wenzhou occupy 70 percent or so of the world market. Ningbo is famous for its clothes industry, which accounts for 12 percent of total clothes production in China (Guo, et al, 2002). Wang's study (2006) shows the industrial cluster evidence in Datang, Wenzhou. This area has become a key sock production centre in China and the world. As Wang (2006, p. 112) claims,

A typical case is Datang, an industrial boomtown in Wenzhou, which churns out more than 8 billion pairs of socks a year, accounting for one-third of the world's supply. Datang's development is a result of the development of agglomerative economies and industrial clusters.

Moreover, Marukawa (2006) argues the agglomeration of the synthetic textile industry in Shaoxing. Ding (2007) refers to the industrial clusters of plastic products and mould manufacturing in Yuyao. In particular, his study highlights the eyewear industrial cluster in Danyang. Sheng and Zheng (2004) summarize the development of various regional industrial clusters in Wenzhou, ranging from production of lighters, leather shoes and pens to printing industries. The Chinese case studies suggest the strong tendency of industrial cluster and rapid development of agglomerative economies. In terms of China, particularly in the eastern provinces like Zhejiang and Jiangsu, it is actually difficult to provide an overview of the prosperous industrial development and economic growth without discussing industrial agglomeration. The positive correlation between industrial cluster and regional economic growth is close. As Kang (2007, p. 1) suggests, "Clusters have been a significant component of the provincial economies of coastal China."

Regional industrial clusters in the eastern regions are largely a product of scale economies and spillover effects. In that sense, increasing return to economic scale is the greatest incentive encouraging the industrial enterprises to concentrate in a few places. The evidence of many Chinese regions indicates that regional industrial clusters are the key to explaining local industrial prosperity and economic growth. Regional industrial agglomeration provides many eastern regions with greater opportunities to develop local industries and participate in international economic cooperation. More significantly, it reinforces the overall development capabilities in these regions.

6.3.1 History and Agglomeration

From my point of view, the historical factors affecting regional economic agglomeration should not be neglected. Historical incidents offer some new ideas for understanding the initial emergence of industrial cluster. For example, the Wenzhou development model has been widely associated with regional industrial cluster. Without question, the fast economic and industrial growth in this place has benefited from industrial agglomeration during the reform period. Significantly, in their book '*Wenhua yu Zhejiang Quyu Jingji Fazhan*' (Culture and Regional Economic Development in Zhejiang), Qi, et al, (2001) point out that Wenzhou has become a famous handicraft and commercial city since the Tang and Song Dynasties. The local people in Wenzhou are traditionally skilled in handicrafts. In its long history, the china, silk, paper, boat and leather products of Wenzhou have enjoyed a very good reputation.

In terms of Liushi, Wenzhou, as previously mentioned, has become the main production centre for the low-voltage electric parts industry in China. Regional agglomeration of this industry is notable in Liushi. However, the research conducted by Shi, et al, (2002) highlights the role played by history in developing the industrial cluster in this small place. Without two peasants named Chen Qingyao and Chen

Weisong, the emergence and further development of the electrical parts industry in Liushi is hard to image. In 1970, when visiting a friend in Anhui Province, Chen Qingyao found that many kinds of small electrical parts were in short supply. The inherited commercial talent of this peasant played a key role in his establishment of the family-based factory producing low-voltage electrical parts. Later, his factory was further expanded. Many local people followed him and engaged in the same kind of industry. Therefore, the emergence of an electrical parts industrial cluster in Liushi naturally occurred. As Marukawa (2006, p. 14) rightly states,

Once the person succeeds in business, the neighborhoods will try to imitate the innovator's success. The industry will spread in the village or township and hence create an industrial cluster.

As far as Shanghai is concerned, the regional concentration of high-tech industrial sectors and foreign-oriented enterprises is largely a result of its glorious trading history. Numerous scholars stress this point. Shanghai was the key sea port in China and worldwide before 1949 (Tregear, 1980). By 1553, Shanghai had already become an important domestic trading port in the Lower Yangtze Delta region (Zhang, Xiong and Shen, 2002). In the early book '*Shanghai: Key to Modern China*', Murphey (1953) points out that Shanghai had not only become the key port in the Yangtze Delta, but also it was a centre of trade between the northern and southern regions of China by the early 1800s. The data presented in Table 6.1 below demonstrates the dominant position of Shanghai among all of the major Chinese ports in terms of both foreign and domestic trade during the period between 1870 and 1920. In 1870, the total value of foreign and domestic trade in Shanghai was 186 million haiguan taels, which was more than three times higher than the second-ranking Chinese port – Guangzhou. Moreover, the gross value of trade in Shanghai was up to 1036 million haiguan taels in 1920; the number was four times higher than the second-ranking port – Tianjin, which was only 225 millions.

*Table 6.1: Foreign and Domestic * Trade in Leading Sea Ports within China, 1870-1920*
 (Gross values of trade, in millions of haiguan taels, landed at or shipped from the leading treaty ports in China in each year listed.)

	1870		1880		1890		1900		1910		1920	
	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic
Shanghai	90	96	88	112	101	136	160	181	316	382	504	532
Canton	28	24	30	26	28	24	34	26	47	38	59	46
Hankow	18	16	21	25	14	42	15	52	13	58	13	73
Tientsin	15	9	16	14	18	25	32	38	64	68	108	117
Tsingtao							5	4	15	9	14	12

Sources: Figures compiled from China: Maritime Customs, Annual Reports and Returns of Trade.

* i.e., the domestic trade which was recorded by the Maritime Customs, only a part of the total volume of domestic trade at each port, or of China as a whole.

Reproduced from Murphey (1953, pp. 121)

Therefore, the significance of historical factors in analyzing the development of regional industrial cluster and economic agglomeration needs to be properly recognized. Without tracking back through history, it is difficult to understand the development of regional industrial agglomeration. This is precisely why Wen's argument (2004, p. 337) is reasonable,

Without retrospect of the development history of Chinese industries and knowledge of regional industry comparative advantage in technology, it is not possible to distinguish the locational difference in this industrial agglomeration.

However significantly, in my opinion, geography is the fundamental cause of industrial agglomeration rather than history. The determining role played by geographical factors in shaping the emergence and development of industrial cluster is notable. In precise terms, locational advantage is the key to development of regional industrial centres. For example, Kunshan is well-known for its regional concentration of Taiwanese IT and other high-tech enterprises. This place is steadily becoming an important manufacturing centre of high-tech products. The study conducted by Lai, Chiu, and Leu (2005) suggests that foreign capital inflow to Kunshan from Taiwanese investors account for one-tenth of total Taiwanese investment in mainland China.

The emergence of the IT industrial cluster in Kunshan is closely associated with its locational advantage of large market proximity (Shanghai and Suzhou). For example, the road distance between Kunshan and Shanghai is only about 50 kilometers. Therefore, the investors have sufficient incentives to set up manufacturing factories in Kunshan. They can reduce production costs and target the large consumer market of Shanghai and other developed regions located within the Yangtze River Delta. Marton's study (2002) suggests the importance of favorable location affecting the industrial cluster in Kunshan. Jiang's study (2003) highlights the importance of the short distance to Shanghai and convenient transport to the development of a manufacturing base in Suzhou and its county-level administrative region Kunshan. Moreover, Wei (2002) also discusses the case of Kunshan, he stresses the advantages of geographical location and large market proximity in this region, and the superior transportation system in connecting Kunshan with the other places in the Yangtze River Delta, such as the water route (Grand Canal), rail route (Nanjing – Shanghai line) and motorway (National Highway 312).

As far as Shanghai is concerned, geography is the key influence on regional economic agglomeration and foreign trade concentration. The historical development of foreign and domestic trade and its important trade position in China are largely shaped by the strategic location. In terms of trade development and commercial prosperity, Murphey (1953) stresses that the distinction of Shanghai among the cities located within the Yangtze Delta lies in its geographical proximity to the East Asian coast, and the proximity to the trade hinterland with convenient water transportation. In the book *Shanghai: From Market Town to Treaty Port, 1074 – 1858*, Johnson (1995) gives some detailed discussion regarding this issue. Why did the British choose Shanghai as one of the Treaty Ports? Why was this place so attractive? It is because Shanghai was already a well-developed coastal

port and commercial centre of the Yangtze Region and China due to its favourable coastal location. Captain Balfour, who was a British consul in Shanghai during the period between 1843 and 1846, pointed out that the selection of Shanghai as a Treaty Port was due to its strategic geographical location in the Yangtze River Delta. By setting up the factories in this place, the British producers could easily access the inland market of China and the British products could be sent to the large consumer market across the whole Yangtze area (Johnson, 1995). As she states,

British officials had now confirmed the reports of missionaries who had managed to visit the city earlier that Shanghai was a pivotal point for trade with the interior and in general could be considered a pleasant place, suitable for foreign residence. (1995, p. 180)

In general, industrial cluster and economic agglomeration mainly occur in coastal areas or in regions inheriting the locational advantage of large market proximity. Cavalcanti, Souza, and Yu's study (2006) suggests that the majority of regional industrial clusters have located within Bohai Bay Rim, Yangtze and Pearl River Deltas.

In the Chinese context, most of the rich and developed cities like Dalian, Qingdao and Xiamen, where many manufacturing industries, economic activities and people are concentrated, have inherited geographical advantages, including both coastal proximity and low transportation costs. Moreover, the economic prosperity of Shanghai has also benefited from the natural geographical advantage (the largest port in China). Due to the reduction of transportation costs in the interior regions through technological innovation and infrastructure improvement, the inherited locational advantages of these coastal areas have become less important and prominent. However, these cities are still achieving economic prosperity and further development; in addition, their leading positions of economic growth seem to be unshaken. Many industrial firms still prefer to locate in these cities rather than relocating to the other peripheral regions. Therefore, there are sufficient reasons to claim that the self-reinforcing pattern of agglomeration effect might strengthen the tendency of industrial and other economic activities to concentrate in the coastal cities; it further reinforces the existing leading economic position of these cities even when the natural geography advantages becomes less significant.

In general, history makes a difference to the formation of regional industrial cluster. The positive correlation between the historical factors and economic agglomeration is evident. As far as the initial development of regional agglomerative economy is concerned, history cannot be ignored. Nevertheless, geography is the key factor in shaping the industrial clusters rather than history. It is logical to expect that the role played by geography is important to an understanding of the economic growth in Guangdong. Therefore, one main hypothesis proposed by this research is to test whether or not these geographical efforts are, indeed, crucial to the regional development in Guangdong.

6.4 The Developing Trend of Regional Inequality

In general, even if economic development reaches a high stage, the evidence of many Chinese regions suggests that the trend of regional income convergence has not occurred to any desirable degree during the reform period. By contrast, the strong tendency of regional divergence is prominent. Numerous previous scholars have done some important research on the issue of intra-regional income inequality within the Chinese provinces, and the income disparity between the coastal core and inland peripheral regions. Many of these studies demonstrate that regional inequalities have been getting worse since 1978 (e.g. Smith, 1991; Wei, 2000; Démurger, 2001; Wang, et al, 2004). For example, Smith (1991) indicates that the core (northeast region) – periphery (western region) pattern of spatial inequality is notable. By adopting the inequality indicator Coefficient of Variation (CV)³² for per capita rural gross social products, Yang (1997) argues that the economic growth gap between the core and peripheral areas within the majority of Chinese provinces had been widening during the 1980s. The research made by Jian, Sachs and Warner (1996) suggests that the coastal-interior pattern of economic inequality had intensified in China during the 1990s.

Similarly, Wang and Hu (1999) point out that although spatial economic disparities narrowed during the 1980s; both the inter-regional, inter-provincial, and intra-provincial levels of economic disparities in China had increased substantially during the 1990s. Démurger (2001) argues that the spatial economic development in China showed a divergent rather than convergent trend during the 1990s. The regional disparity between the developed coastal and least developed interior provinces was intensified. Natrajan's study (2006) supports the arguments made by Démurger, Wang and Hu; suggesting that inter-provincial economic disparities were narrowed dramatically between 1980 and 1985. Meanwhile, accompanying rapid economic growth, Natrajan (2006) argues that regional income distribution had become quite unequal and continually worsened during the period from 1980 to 2005. His empirical research shows that Gini coefficient was up to 0.45 in 2005 from only 0.33 in 1980, which was an increase of more than 36 percent during this period. Martin and Ottaviano (2001) claim that China has been facing increasing spatial inequality between the coastal and interior regions; the geographical advantage of coastal regions, with their huge inflow of foreign capital, encourages industrial clustering. Therefore, the leading economic position of these areas is further strengthened: meanwhile, the interior regions are far behind the coastal regions in terms of economic growth. A report by the UNDP (2002) discusses the widening spatial inequality between the rich coastal eastern and poor inland western regions. As the UNDP (2002, p. 45) suggests,

... China used to be a poor but by international standards relatively equitable society. However, many inequalities do exist, and these have grown substantially since the reforms took off in the early 1980s.

For example, the UNDP report (2002) points out the sharp social and economic gaps in terms of income per capita, education, and life expectancy between Xizang and Shanghai in 1999.

³² CV is the abbreviation of one inequality indicator: coefficient of variation.

Moreover, as far as the intra-provincial economic disparities are concerned, the core – periphery pattern of regional divergence can be severe. Scholars have provided numerous studies based on Chinese provinces (e.g. Chen, 1998; Lyons, 2000; Wei, 2000; Wei and Ye, 2004). In terms of Tianjin, Chen's study (1998) demonstrates that Gini coefficient of per capita income distribution had increased to 0.365 in 1995, up by 38 percent from 0.2635 in 1981. Thomas Lyons has done detailed research on Fujian: which was another of the earliest Chinese provinces to implement the 'reform and opening' policy. Although this place has achieved generally rapid economic and industrial growth during the reform decades, Lyons's study (2000) highlights that Fujian had been facing the challenge of increasing inter-regional economic disparities during the same period. He argues that rapid economic development has mainly occurred in the northern and southern regions from Longyan to Xiamen (in terms of output per capita); he referred to it as the "north-south belt"; while the eastern and western regions have suffered from slow economic growth, and were behind other regions in Fujian. Lyons's CV analysis of output per capita shows that the regional disparities between total 68 counties had risen dramatically between 1978 and 1995 (0.714 in 1995 vis-à-vis 0.565 in 1982). As Lyons (2000) points out,

With the reforms and opening up of the 1980s and 1990s, Fujian experienced a phenomenal burst of economic growth. ... One cannot expect growth at this breakneck pace to spread evenly across every corner of the province. Rather, one might expect an initial widening of interregional disparities ... (Cited in Yeung and Chu, 2000, p. 328)

Yang and Zhang's empirical finding (2001) shows that the rural economic inequality was higher than in the urban areas in Zhejiang Province, for example, in terms of the Gini coefficient of urban household income, the urban inequality was 0.231 while the equivalent figure in rural areas was 0.341 in 1999. Wei and Ye's analysis of regional development disparities within Zhejiang (2004) suggests that both the inter-municipal and inter-county levels of regional economic disparities have increased considerably since 1978. Moreover, in their later study, Ye and Wei (2005) further highlight that the regional economic disparities within Zhejiang widened during the reform period mainly due to regional industrial agglomeration and geographical location.

Jiangsu is another eastern province which has enjoyed numerous state-oriented preferential development policies and achieved dramatic economic development over the past two decades. However, Jiangsu Province has also been facing the problem of marked uneven regional development and widening spatial income disparities during this period. For example, in terms of industrial and economic development, Honig (1992) claims that the sharp difference between Sunan and Subei regions. Rozelle (1994) argues that the income distribution among counties within Jiangsu has become extremely uneven; for example, in terms of the inter-county level of inequality, the Gini coefficient for household income per capita was 0.44 in 1988/1989. Wei and Fan's empirical research (2000) points out that the interregional economic disparities within Jiangsu had been intensified during the reform period. Wei Yehua is one of the most representative scholars studying the regional inequality within Jiangsu. Wei (2000) highlights that a scale

of worsening spatial disparity was found in Jiangsu, from the rich and developed Sunan region, the medium-developed Suzhong region, to the poor and under-developed Subei region. Wei (2000) found that inter-regional economic disparity had increased only a little between 1978 and early 1980s due to the successful rural reforms and rapid growth of agriculture and other farming sectors; but the gaps of uneven spatial development (Sunan – Suzhong – Subei) have been getting larger since the middle the 1980s (the CV of inter-regional disparity was 0.64 in 1995; up by 200 percent compared to the equivalent figure in 1978). In terms of provincial comparison between Jiangsu and Guangdong, during the early 1990s, the interregional income inequalities within Jiangsu were more serious than in Guangdong. For example, in 1995, the CV figures were 0.64 and 0.559 in Jiangsu and Guangdong respectively. By contrast, Wei believes that the inter-county economic disparities in Jiangsu decreased due to the under-development of the old cities. Wei stresses that,

During the reform period, interregional and rural inequalities have increased further, and the gap between Sunan and Subei has been intensified, as richer counties in Sunan registered more rapid growth than poorer counties in Subei. (2000, p. 167)

Hendrischke's study (1997) suggests regional development inequality between the south-east and north-west regions in Guangxi. The economic growth in the south-east region is generally faster than the other regions in Guangxi; but the north-west region has fallen behind; this can be demonstrated from the significant number of people living in poverty in this region. There is a total of around eight million people residing in the north-west part of Guangxi and Hendrischke found that most of them were still living under both the national and international poverty lines. Neither can Shandong Province be exempted from the problem of worsening spatial economic disparities between various regions during the reform period. Chung's study (1997) demonstrates that the economic growth and wealth of Shandong have been mainly concentrated in Jiaodong Peninsula (*Jiaodong bandao*); this is especially true of Qingdao, Yantai and Weihai municipalities. However, the interior and western regions have developed slowly and lagged behind the other regions.

In general, the previous studies suggest that the regional economic disparities in China have intensified since the early 1990s; this was particularly prominent after 1992, at both the inter-provincial and intra-provincial levels. The UNDP (1999) has also made similar suggestions. Moreover, many scholars share their concern over the challenge of worsening regional development inequality within China during the 'reform and opening' period (e.g. Yang, 1997; Wang and Hu, 1999; Wei, 2000). They argue that the government must address this issue seriously and take strong action to resolve the regional inequality problem in China; otherwise it would not only affect social harmony and long-run economic growth, but also threaten the political stability and national utility of China.

The regional income comparison between the coastal and inland provinces demonstrates a divergent trend; moreover, the income gap between the wealthy core and poor periphery within the developed and under-developed provinces reflect an even more divergent tendency. The regional income comparison between the

selected provinces Guangdong, Shandong, Henan and Guizhou during the reform period, which is presented in the following chapter, clearly supports this argument. Regional inequality measurement demonstrates that both the CV and Gini coefficient had increased substantially in all of these four provinces during the reform period. Inequality measurement analysis suggests that spatial income disparities tend to be severe in both the economically rich and poor provinces. Both the developed and least developed provinces face the challenge of rising intra-regional disparities; and the personal income distribution is unequal between core and peripheral regions like Shandong and Guizhou. For example, few regions located in Guizhou are wealthy: like the county regions surrounding Guiyang and Zhunyi municipalities. The average income growth in these rich regions has been impressive since 1980. In terms of GDP per capita, this study found that the absolute income level in these few regions in Guizhou is not much lower than the rich regions in Guangdong. However, the majority of regions are extremely backward and under-developed; this is especially true of the north-west and south-west ethnic minority regions.

It is important to note the following points. The widening regional economic inequality in China does not support the U-shaped regional inequality theory proposed by Kuznets. Moreover, the regional income convergence claimed by the neoclassical economists has been largely rejected by the Chinese evidence. The income inequality between the various regions within China demonstrates a divergent trend rather than a convergent one. Regional income disparity could be high even if the local economy of a place enters the mature development stage. The widening tendency of regional income divergence is notable in the developed areas as well as in the least developed areas. Almost all of the Chinese provinces have faced the problem of unbalanced regional economic development and intensifying income inequality between the core and peripheral regions during the reform period. The core – periphery pattern of spatial economic disparity is the main development theme dominating the regional income inequality in China during the reform period. Skinner's study supports this argument. This is what Skinner (1994, p. 49) claims,

But the major fault line in the geography of China's development is not that of coast as against interior. It is core versus periphery within China's macroregional economies.

Inter-regional disparity has risen in the other regions of China. The regional evidence in China has not supported the theories suggested by Kuznets and neoclassical economists. Therefore, it is reasonable to hypothesize that the spatial inequality has increased in Guangdong too. It is logical to predict a divergent trend of spatial inequality in Guangdong. Detailed study with regard to the development tendency of regional inequality in Guangdong will be demonstrated in Chapter Nine.

6.5 Skinner's Macroregion Model

In his edited book *The City in Late Imperial China*, Skinner (1977) proposes a new regional theory of China – the macroregion model³³. Based on spatial variation

³³ "Traditional regional geographies of China, by both Western and Chinese scholars, have

in terms of economic development, sociopolitical, agricultural and demographic factors, he separates China into nine macroregions for geographical analysis; these nine macroregions include North China, North-west China, South-east Coast, Lingnan, Upper Yangtze, Middle Yangtze, Lower Yangtze, Yun-Kwei and Manchuria. Skinner (1977) argues that the cities located in each of these nine regions have much closer economic and social links with the other places within the same macroregion than with the other macroregions.

However, Skinner's macroregional geographical analysis has faced numerous criticisms from the academic community. For example, Cartier (2002) argues that the macroregion model has not involved as much study of the regional geography and landscape in China as Skinner claims. In addition, Skinner's macroregion theory has not given enough consideration to regional differences, the unique characteristics of each region and its individual local human history.

To my understanding, one of the important justifications and theoretical bases for the macroregion theory proposed by Skinner is the difficulties for interregional economic and trade communication and the high transportation costs caused by geographical distance. Nevertheless, Skinner ignores the role of dramatic development of the interregional transportation facilities in increasing regional communication and substantially reducing the interregional transportation costs. Therefore, the arguments suggested by Skinner's macroregion model should be reconsidered in the present analysis of regional geography. Cartier (2002, p. 117-18) suggests that,

Yet the spatial structure embedded in the macroregion perspective discourages investigations into process of economic activity and interrelations between different regional formations and interrelated spatial processes, such as activities that cross regional boundaries in long-distance and maritime trade.

The inter-regional trade evidence of many individual commodities illustrates that the economic and trade communication between various Chinese macroregions is notably close and active. Geographical distance and high transportation costs have not negatively affected trading between the various macroregions. Even when facing extreme geography and transportation difficulty, the cross-regional movement of goods and people has developed substantially. Moreover, inter-regional transportation improvement accounts for prosperous inter-regional trade. This issue simply cannot be eliminated. Chinese cases seem not to support the macroregion model developed by Skinner and lay its credibility open to question. For example, salt production and domestic trade in China discussed in Chiang's research (1983), suggest close interregional economic and trading cooperation between the different macroregions within China. Chiang considers that as salt is one of the staple goods for the people, analysis of the salt trade is sufficiently representative of the general situation of interregional trade and economic interaction between the various macroregions. Owing to the sporadic distribution of salt production, Chiang (1983) points out that the interregional salt trade is prosperous. According to the data given by Chiang (1983, p. 198), Huainan and

assessed its characteristics in the context of natural regions or provinces." (Cartier, 2002, p. 83)

Liangche are the two main salt producers in coastal China, and account for 18.9 and 17.1 percent of the total domestic salt production. However, in a few regions like Shenkan, the percentage share of salt production in China was notably small. Under these circumstances, the interregional salt trade is crucial. Salt transportation between the different coastal macroregions is quite convenient and water transport costs are low. For example, the salt produced in Huainan, Jiangsu, can be delivered to the consumer market up to 1,200 kilometers away. By contrast, road is the main means of transporting the salt produced in the inland areas. The privileged Chinese merchants were heavily involved in the interregional salt trade.

Moreover, the prosperous rice trade and movement of rice between the different macroregions are also discussed by Chiang. For example, from the twelfth century, the rice grown in Hunan, Hubei and Jiangxi was shipped downstream to the regions located in the Lower Yangtze River Delta (Chiang, 1983). Similarly, Chuan and Kraus's study (1975, cited in Sands and Myers, 1986) suggests that the provinces located in Lower Yangtze River Delta had imported large amounts of grain from other macroregions in China. Sands and Myers (1986) claim that the inter-regional grain trade has played a key role in the survival and development of the individual macroregions. Moreover, the merchants traded cotton cloth, originally produced in Suzhou and other Jiangnan regions, to the other macroregions. Therefore, the cotton industry expanded and further developed (Nishijima 1966, cited in Sands and Myers, 1986). In fact, since the sixteenth and seventeenth centuries, the domestic merchants have participated in these prosperous economic activities. More significantly, the merchants not only actively promoted inter-regional trade, but also they were involved in trade with other nations. For example, Atwell (1982) provides detailed research of the international silver trade between China and Japan.

Sands and Myers (1986) have made strong criticisms of the macroregion model proposed by Skinner. They suggest that not every individual macroregion has achieved such strong and independent spatial characteristics in terms of grain and price correlation. By contrast, economic communication and trade interaction were extremely close among the various macroregions: rather than the casual and weak inter-regional links suggested by Skinner. Sands and Myers (1986, p. 737) point out that,

We were concerned with whether Skinner's concepts of macroregion, core and periphery were defined with sufficient rigor and were sufficiently clear and unambiguous to have explanatory power. Our conclusions are the following: his concepts are difficult to test, they are seriously flawed, and they lack real explanatory power.

Sands and Myers's criticism is largely shared by Lavelly (1989). Lavelly maintains that Skinner's macroregional model is difficult to analyze. The internal functioning mechanism of this model remains unknown. He believes that more supporting statistical figures need to be collected and new empirical tests should be created.

Numerous studies suggest that geographical distance has not interrupted the economic ties and inter-regional trade between the various macroregions. Inter-regional trade and economic interaction flourished in Ancient China.

Therefore, when defending his macroregion model, Skinner's key argument of difficult inter-regional trade caused by high transportation costs and long distance seems to be weak and inappropriate. The individual macroregions are not so unique and distinguished that they can be virtually separated from each other. The interregional interaction between the macroregions is close in terms of trade, economic activities and also people. With further inter-regional infrastructure improvement and transportation cost reduction in contemporary China, it is logical to expect that the inter-regional economic links will become stronger. Skinner's macroregional theory is very problematic.

In my view, traditional provincial boundaries are still important to the study of regional economic geography in China. There are two main reasons. First, owing to the political system and administration practices, the provincial institutions have strong influence within their own administrative boundaries (for example, prefecture, city or county) in terms of economic development and policy planning. Secondly, in many Chinese cases, intra-provincial communication and linkages are far stronger than inter-regional interaction in the field of economy, society, trade and culture. Within Guangdong, the regional linkage between various places is close, but the same is not true of links to the outside. For example, the most remote and border areas like Chaozhou and Zhanjiang maintain varying levels of economic and trade communication with the neighbouring regions outside Guangdong, like Fujian and Guangxi, but the economic linkages between Chaozhou, Zhanjiang and the provincial centres are significant. The cross-provincial macroregion theory has not seriously addressed political factors or the issue of close intra-provincial communication. As far as regional differences and special spatial characteristics are concerned, dramatic provincial variation has not been taken into account in macroregional geographical analysis. Therefore, some arguments suggested by Cartier are inappropriate. The traditional provincial approach to understanding regional economic geography is still useful and relevant. Cartier is over-confident about the general credibility of the macroregional analysis; she seems to overestimate the theoretical development of macroregional geography. As Cartier (2002, p. 128) alleges,

Regional analysis is no longer necessarily based on the bounded territorial region, and questions have shifted to concentrate on process of regional formation, regional rise and decline, and comparative regional analysis.

6.6 Conclusion

As far as China is concerned, one of the important causes of the increasing spatial economic disparities between the eastern, central and western regions is location and other topographical features. The western region, especially the mountainous area, has been facing extremely poor geographical and climatic disadvantages. Locational disadvantage and high transportation costs have negatively affected economic growth in the under-developed regions. In addition, the extreme weather has caused serious challenges for farming and industrial development in the inland western regions. However, the inherited advantages of location and other

geographical aspects have boosted local industrial and economic growth in the coastal eastern regions.

In terms of the trend of regional inequality, the regional economic development within China varies considerably. The spatial income gap between the developed core and under-developed peripheral regions is vast. The severe regional economic variation is striking. For example, economic growth in Shanghai (as shown in the average growth rate of GDP per capita) is dramatically faster than in Guizhou and Qinghai. Moreover, macroregional comparison reveals notable differences between the prosperous eastern and impoverished western regions. In general, the key messages generated from study of Chinese cases are as following. First, regional income divergence is not only shown in the coastal – inland province pattern; but also more significantly, in the wealthy core – poor periphery pattern within the developed and least developed provinces. Secondly, the developing tendency of regional income inequality is divergence rather than convergence. Regional income convergence suggested by the neoclassical growth theory has not occurred to any appreciable extent. Neither has the inverted U-curve hypothesis of regional income inequality proposed by Kuznets been supported by Chinese evidence with regard to unbalanced regional development during the reform period.

Lastly, Skinner's macroregional model enlarges our understanding of regional economic geography in China. To some extent, the concept of macroregion is useful for assessing regional economic and trading interaction. His model indicates that the traditional provincial approach in addressing the inter-regional economic communication has some limitations. Nevertheless, Skinner's theory suffers from some serious problems; his argument is not completely convincing. The major argument on the inter-regional trade difficulty caused by high transportation costs and long road distance suggested by Skinner is not justified. Geographical distance in association with high transportation cost has not entirely blocked the economic ties and inter-regional trade activities between the different macroregions in China. In fact, as shown by the Chinese evidence presented earlier, inter-regional trade and commerce have been steadily developed since at least the days of the Tang Dynasty. Regional commerce and economic linkages are active and close. The influential merchants have actively participated in this process. The individual macroregions are not so spatially independent that they can be totally distinguished from each other.

The Chinese evidence suggests that location and transportation costs make a difference to economic growth; coastal location and low transportation costs have contributed to the economic prosperity in the eastern region. More significantly, industrial clustering and agglomerative economies are crucial to understanding of regional economic development and industrial expansion. As far as the changing trend of regional inequalities within China is concerned, a divergent trend between the various regions is notable. Therefore, it is reasonable to expect that the coastal regions within Guangdong with good and fast access to large markets may have achieved rapid economic growth. Moreover, a generally widening core-periphery pattern of regional disparity could be predicted to be the central development theme in Guangdong. A detailed discussion of Guangdong is presented in Chapter Nine, Ten and Eleven. The next chapter presents fieldwork design and methodology adopted by this research for the study of Guangdong.

Chapter 7 Methodology and Fieldwork Design

7.1 Introduction

This chapter describes fieldwork research design and research methodology adopted in this study. The methodological includes both field interviews and official data analysis. By applying a combination of methods of face-to-face interviews and data analysis, I intend to address and test the two main hypotheses proposed by this research. Analysis of research method, particularly the interview method used by this study, will be demonstrated in this chapter. Moreover, the issues of research sample, research reliability and validity and ethical issues regarding this study also discussed here. My personal perceptions and experiences of the field trip to Guangdong during the autumn of 2006 are also presented. In particular, the specific actions taken to ensure the quality of data collected from field research interviews have been highlighted. The final conclusion is drawn in Section Three.

7.2 Research Methodology and Fieldwork Study Design

7.2.1 Research Methods: Questionnaire and Interview

7.2.1.1 Criticism of Questionnaires as a Research Method

Numerous scholars (e.g. Hughes, 2002; Saunders, Lewis and Thornhill, 2003; Bryman, 2004) have done important research on the issue of the questionnaire research method. They conclude that questionnaires can suffer from many limitations and problems. For example, they are unable to fully interpret the opinions of people, and the researcher has almost no opportunity to help the participants in understanding the precise meaning of research questions mainly due to the absence of face-to-face communication and visual interaction. One of the biggest issues with questionnaires is probably that reading, understanding and answering the questions entirely depends on the participants themselves, whereas interviewers can, during the interview conversation, provide assistance to the interviewees in their understanding of complex words and concepts involved. Hopkins, Stanley and Hopkins (1990, p. 50) claim that, "most people are not very thorough in a research sense and this is all the more true about dealing with questionnaires – an activity which typically they do not enjoy or benefit from in any way." Moreover, the questionnaire can be subject to shortcomings in credibility and reliability as difficult questions using complex words might be left unanswered by the participants. In addition, the researcher lacks the power to identify who actually answers the questions of the questionnaire. The people who completed the questionnaire might not be the potential targeted participants. Bryman's study (2004) clearly argues this point. In highlighting the limitations of the questionnaire research method, he argues that incomplete data is a result of questions being ignored by interviewees.

In addition, a researcher relying on questionnaires lacks the opportunity to ask further questions in order to double-check the validity of the data and thoroughly address the research topics.

7.2.1.2 Advantages of Interview Method

In contrast to questionnaires, the interview has numerous advantages in conducting academic research. First, the interview allows the researcher to explore the social issues and people's own perceptions in depth by encouraging the expression of interviewees' personal feeling and opinions. As Gorden (1987, p. 11) highlights,

Interviewing is most valuable when we are interested in knowing people's beliefs, attitudes, values, knowledge ...

In my experience, the most in-depth, meaningful and rich research data and other relevant information came from the one-to-one interview approach, which allowed me to measure and analyze regional inequality and its main causes within Guangdong and make meaningful comparisons of inter-regional economic inequality between the provinces within China. More significantly, the sort of data and information collected through the face-to-face interview are generally more reliable and accurate than data from postal questionnaires. In my belief, sending a questionnaire by email or post to China would show much less respect to the potential participants than personal interviews. Many interview participants strongly share this view. Secondly, people are more interested in participating in the face-to-face interview rather than in the self-administrated questionnaire, even if the interview takes more time due to the natural characteristics of human beings. People want to be heard and listened to when they are expressing their own opinions and feelings; in addition, people like to share their opinions with other people. The interview method provides exactly this opportunity. Gillham's study (2000) offers a detailed discussion regarding this issue.

Thirdly, the total interviewee sample is relatively small and easily managed; therefore, an in-depth investigation of people's opinions becomes possible and meaningful. The questions which have not been answered by the interviewees can be asked again during time allocated to free conversation and open discussion. By contrast, it is almost impossible to achieve this goal by using questionnaires. Gillham's study (2000) has suggested the essential criteria for using the interview research method (see Table 7.1 below).

Table 7.1: Criteria for Conducting Interview Research

➤ Small numbers of people are involved
➤ People are accessible
➤ Most of the questions are 'open' and require an extended response with prompts and probes
➤ Everyone is 'key' and you can't afford to lose any
➤ The material is sensitive in character so that trust is involved
➤ Anonymity is not an issue, though confidentiality may be
➤ Depth of meaning is central, with only limited approximation to typicality
➤ Research aims mainly require insight and understanding

Reproduced from Gillham (2000, pp. 11)

Interview has been a very popular method of data collection in field research in China since the late 1970s (O'Brien, 2006). As we understand, interviewing exiles from mainland China who were working and living in Hong Kong had been a common focus of research during the 1970s. This research practice changed substantially during the reform period. In this study, semi-structured interviews

were conducted during the field research period of autumn 2006. Both open-ended and closed-ended questions were asked in order to collect detailed and useful information. The semi-structured interview is widely argued to be the fullest approach to the 'why' kind of investigation of a research topic. (Saunders, Lewis and Thornhill, 2003)

However, it is important to note that the interview method is not without its critics. Indeed, the interview has its own limitations. Hughes (2002) believes that certain factors might affect the quality of interview data, such as cultural differences of interviewees, small interview sample size and over-dependence on the interviewer's and interviewees' own personalities to provide accurate information. The most important issues are probably time and cost considerations. For instance, Bryman (2004) points out that the researcher needs to take longer time and spend more money on an interview than a questionnaire; this is particularly true of the interviews with large sample size. Moreover, Bryman (2004) suggests that the interviewing method might suffer from problems of being too subjective. This is because the interview data was obtained from private conversation; it also due to over-dependence on the interviewer's own judgment, efforts, and opinions in selecting, interpreting and analyzing the information. Moreover, relative to the questionnaire, the interview data might lack sample representativeness due to relatively small numbers of interview participants. As Bryman (2004, p. 284-85) claims,

When participant observation is used or when unstructured interviews are conducted with a small number of individuals in a certain organization or locality, [quantitative researchers] they argue that it is impossible to know how the findings can be generalized to other settings. (Bold content added)

In my view, it is crucial to provide assistance when necessary and to explain to the participants any complex concepts and difficult words in the research questions. My opinion is based on two main concerns. On the one hand, over-abstract topics are unlikely to stimulate the interests of the Cantonese or encourage them to actively participate in open discussion. In his book '*Qishi ni budong Guangdongren*' (In fact: You Don't Understand the Cantonese), Ye Shuming (2005) has provided some detailed discussion on this issue. On the other hand, more significantly, it is important to note that there are still quite large numbers of people who are either poorly educated or completely illiterate: this is especially true of elderly local residents of the poorest and least developed regions within Guangdong. However, these people simply cannot be eliminated from our potential interviewee sample, as they might have lots of personal experiences and rich information regarding the topic of spatial inequality. In terms of the total illiterate population aged 15 and above, and their share percentage of total population aged 15 and above, the figures presented in Table 7.2 illustrate that the level of illiteracy in Guangdong is unacceptable. In fact, Guangdong has still a long way to go in order to develop basic education and reduce its total number of illiterate population: although the illiteracy level is still much lower than in many western and interior regions of China. As far as Guangdong is concerned, the share percentage of illiterate population to total population aged 15 and above was still up to 6 in 2005. This

illiteracy figure is much higher than the equivalent figures in Beijing, Tianjin and Shanghai.

Table 7.2: Illiterate Population Aged 15 and above by Sex and Region within China (2005) (unit: person)

Region	Illiterate population	Male	Female	Illiterate population to total aged 15 and over (%)	Male (%)	Female (%)
Beijing	7152	1393	5759	3.92	1	6
Liaoning	22839	5700	17139	4.77	2	7.1
Tianjin	5793	1303	4490	4.8	2	7.4
Shanghai	11213	2019	9194	5.24	1	8
Shanxi	19472	5673	13799	5.57	3	7.9
Jilin	17994	5436	12558	5.85	3	8.2
Guangdong	57370	10634	46735	6	2	9
Heilongjiang	26633	7584	19049	6.18	3	8.9
Hebei	53638	13842	39797	7.18	3	10.5
Xinjiang	16919	6370	10549	8.32	6	10.4
Hunan	58377	15200	43177	8.58	4	12
Guangxi	40625	9040	31584	8.64	3	13
Hainan	8162	1794	6367	9.76	4.2	15.6
Henan	95891	26929	68962	9.79	5.5	13.9
Jiangsu	83702	17528	66174	10.02	4.3	15.2
Shaanxi	40903	12024	28879	10.33	6	14
Jiangxi	45038	10308	34730	10.54	4.9	15.9
Inner Mongolia	29476	8751	20725	11.25	6.5	16.1
Chongqing	34244	9658	24586	11.65	6.6	16.5
Zhejiang	65172	16687	48485	11.95	6	17.8
Hubei	74114	17902	56212	12.09	5.8	18.2
Shandong	127449	31741	95707	12.38	6.2	18.3
Fujian	49279	10028	39251	12.92	5.3	20
Sichuan	141104	40946	100158	16.61	9	23
Ningxia	10935	3271	7664	18.71	11	26.1
Anhui	119875	34197	85678	19.24	11.2	26.8
Yunnan	89713	27653	62060	20.07	12	28.4
Gansu	54780	16601	38180	20.83	12.7	28.8
Guizhou	75750	19190	56560	21.41	10.7	32.3
Qinghai	13143	4101	9042	24.07	15	33.1
Xizang	11953	4374	7579	44.84	33.4	55.7
National average	1508706	397877	1110828	11.04	5.8	16.1

Note:

1. Data in this table are obtained from the 1% Population Sample Survey in 2005. The sampling fraction is 1.325%;
2. Illiterate population in this table refers to the population aged 15 and above, who are unable to read or have great difficulty.

NBS (2006, pp. 114)

Moreover, as far as illiteracy within Guangdong is concerned, the figures demonstrated in Table 7.3 show that there is considerable regional variation. In general, the share percentage of illiterate population to total population aged 15 and

over in the developed and economically rich regions is significantly lower than in the underdeveloped and poor regions. For example, the share percentage of illiterate population in Shenzhen was only 1.24 percent; however, the equivalent figure in Meizhou was up to 7.66. The illiteracy percentage in Meizhou was not only more than 6 times higher than in the developed regions, but also considerably higher than the provincial average level.

Table 7.3: Illiterate Population Aged 15 and above by Sex and Municipalities within Guangdong (2000) (unit: percentage)

Region	Illiterate population to total aged 15 and over	Male	Female
Shenzhen	1.24	0.5	1.9
Dongguan	1.46	0.5	2.2
Guangzhou	2.8	0.6	5.1
Zhongshan	3.03	0.9	5
Foshan	3.68	0.9	6.6
Jieyang	4	1.6	6.3
Zhuhai	4.06	1.5	6.4
Jiangmen	5.15	1.4	8.7
Shantou	5.17	1.8	8.4
Chaozhou	5.63	1.7	9.5
Huizhou	5.81	2	9.5
Zhaoqing	6.82	1.3	12.1
Yunfu	7.31	1.7	12.8
Meizhou	7.66	1.5	13.5
Yangjiang	8.46	3.2	13.9
Shaoguan	8.52	2.2	15
Zhanjiang	8.74	4	13.6
Heyuan	9.05	2.7	15.2
Qingyuan	9.43	2.9	16.1
Maoming	10.06	4.1	16.1
Shanwei	11.06	4.4	17.8
Provincial average	5.17	1.7	8.6

Note: The population who are aged 15 and above, and are not able or find it very difficult to read are regarded as the illiterate population in this table.

Population Census Office of Guangdong Province (2002, pp. 272-79)

During the field research trip, I found that the educational background of some local people is poor, their literacy levels and reading skills are also relatively low. The Vice-Governor of Guangdong also admits that, “about 40 percent of the residents’ net incomes are lower than 1,500 *yuan* per year in more than 4,000 administrative villages within Guangdong; almost all of these villages are located in the periphery. Owing to high tuition fees and other education costs, many people cannot afford to send their children to school” (Field Interview 1, September 2006). Under the circumstances, it is not difficult to understand that there are numbers of Cantonese who suffer from reading difficulties. Therefore, collection of reliable and accurate data from Guangdong by administering questionnaires is not feasible.

From my point of view, interviews are essential to careful measurement and critical examination of regional developmental inequality within Guangdong during the reform period. On the one hand, the secondary analysis of official statistical data is an important aspect of research; however, it cannot tell us the whole story or reveal the specific picture of regional inequality from the angle of personal perception. In addition, the official statistical data has certain bias and shortcomings due to various difficulties in collecting accurate data and concerns over political stability and one-party control in China. Therefore, obtaining more detailed information and specific data about the research topic requires adoption of the interview method. Thøgersen (2006, p. 190) states that,

It should be kept in mind, however, that local documentary sources in most cases reflect the views of what we in earlier chapters have called 'official China', and they cannot stand alone if we want to understand the mental and social world of ordinary Chinese citizens.

The personal understanding, feeling and opinions of the selected participating interviewees are crucial to investigation of the research topics; as residents living and working in local regions have first-hand local information and opinions to express regarding regional inequality. The issues involved in regional disparity within Guangdong: such as income and daily expenditure, have directly affected their living conditions. The voice and opinions of the masses of ordinary people should be heard; I believe that field research interviews provide a free discussion forum, supplying local people with a good opportunity to freely express their real opinions. This is especially true of the many poor local residents living in the least developed hill regions; these people and their local communities have been ignored and forgotten by central and provincial governments of China.

7.2.1.3 Elite Interview

Besides the interviews with selected ordinary local participants, I have done several special interviews with targeted people from both government and academic institutions in Guangdong. This kind of interview is regarded as the 'elite interview' by scholars (e.g. Gillham, 2000). Gillham further argues that these special people have some kind of superior knowledge and are able to access privy information which is denied to ordinary people; and their particular opinions deserve to be explored.

Certainly, some important general information and data cannot be directly obtained from ordinary local participants; such as detailed background information regarding regional inequality within Guangdong, state-oriented regional development policies, official documentation and other important statistics. Therefore, conducting the elite interview becomes essential. The so-called 'elite interviewees' are the people who are influential, control important policymaking information, and might have been personally involved in policymaking decisions dealing with regional inequality; to some extent, their own perceptions and opinions are important for helping me to analyze the research questions, this is especially true of background knowledge of the regional disparities within Guangdong. Gorden (1987, p. 182) stresses that,

Often the high-status respondents in any organization or community are

much better than the lower-status person in providing an overall view of the organization or its relationship to other organizations. ... They often are more able to report on the past history and future plans of the organization.

In terms of my personal experience during field research, I had some difficulty in accessing and interviewing governmental officers. Indeed, time limitation and interview place restriction made the interview arrangements very difficult to organize. Due to their busy working schedules and egotistic personalities, the government officers are often reluctant to take part in interviews. The research done by Saunders, Lewis and Thornhill (2003) has provided a detailed discussion regarding the challenges of conducting interviews with officials.

During the field research in Guangdong, gaining access to these governmental officers was one of the challenges I faced. It took quite a long time to organize interviews, as they had to be arranged through intermediary people (so-called gatekeepers³⁴) and personal relationship networks (*guanxi wang*). The government officers felt that they had no duty to talk with a mere academic researcher; in particular on such a negative topic and dark side of Guangdong – regional inequality. Academic research was very low on their list of priorities. In addition, I had to convince the officials that I personally was above suspicion; only then were they willing to speak with me and freely express their opinions.

However, the data and information collected from the elite interview should be carefully examined. To be frank, I have some reservations on the quality of elite interview data. It is impossible to guarantee that these influential people in China, who have first-hand knowledge and personal experience about regional inequality, would honestly and freely provide the necessary data and information. They might not want to discuss their own experience or provide detailed information with regard to the state-oriented policies and their decision processes for reasons of self-censorship, departmental regulation and personal suspicion of the interviewer. Bryman's study (2004) has provided discussion on the reliability and objectivity of official data, documentation and information.

In the context of China, collecting data from elite interviews could be particularly problematic. The Chinese political system is authoritarian and less than transparent; the government officers have absolute power without effective 'check and balance' supervisory systems, and the independent media and other non-governmental organizations (NGOs) have also been virtually not-existent in China. As Bond (1991, p. 85-86) argues, "Within a hierarchy these authorities are less subject to supervisory checks and balances than are those in a more democratic political tradition." The study of administrative reform in China conducted by Ma (1999) has also shared this view. Ma argues that the administration system of this nation has been facing the problem of insufficient checks and balances. Power is too much in the hands of the Communist Party, which has not been effectively checked and balanced by any third independent body. Moreover, Zou's empirical research (2000) offers some evidence to demonstrate the negative consequence of this absence of

³⁴ According to the definition given by Miller and Bell (2005, p. 55), gatekeeper are "those who are in a position to 'permit' access to others for the purpose of interviewing."

proper supervision of power. Since 1993, there have been an increasing number of investigatory cases by disciplinary organizations in China (annual growth nine percent); in addition, punishment of officials has also shown a fast growing trend (annual growth 12 percent).

Therefore, it is reasonable to claim that in order to maintain the power of the Chinese Communist Party (CCP), the officials might deliberately give inaccurate and non-objective information and data during elite interview conversations. Moreover, the interview was concentrating on discussion of the dark and negative side of Guangdong. Although the issue of regional inequality is no longer a very sensitive or forbidden topic in China, it is apparently a subject that not all government officers would wish to discuss in depth with an interviewer. Owing to the virtual non-existence of civil society and the tight governmental controls, would the officials feel comfortable taking part in academic research and frankly expressing their opinions? In an early study, Bond (1991) points out that the challenges in conducting field research in China are mainly due to the less than transparent political system.

Moreover, the data obtained from the elite interview might also have been subjected to various bias and omissions due to considerations of political stability and power control by the officials. Bryman's study (2004) has clearly highlighted this point. Therefore, in order to obtain more objective data and try to eliminate bias content in the elite interview as much as possible, in terms of this study, any data and content collected from the elite interview has been examined and strictly checked through various channels, such as review of local written information, personal judgment and field interviews with ordinary people and academic scholars.

7.2.2 Research Sample

Mason (2002) and Bryman (2004) have done important studies on research sampling. They argue that while there is no ideal sample size, it should be sufficient to generate meaningful research and to analyze the research questions and hypotheses. There are not any fixed samples size can be told in advance.

My field research sample is separated into two groups: the ordinary and elite sample interviewees. On the one hand, there are nine selected elite interviewees participating in the elite interviews, representing different government and research institutions; they are either the leaders of individual government departments or academic scholars. On the other hand, 15 local interviewees have been successfully interviewed in each of three targeted counties and districts within Guangdong. A total of 45 local interviewees participated in field interviews. The interviewees' main status and interview questions are presented in Appendix 8. These 45 people are chosen from three regions in Guangdong – Meixian, Wuhua and District of Guangzhou – which represent the different developmental levels of regions in terms of average per capita GDP, including the developed, middle and under-developed counties. According to the data presented in Table 7.4 below, relative to the provincial average of per capita GDP, District of Guangzhou had obtained much higher average per capita GDP during the ten years from 1995 to 2005; more than three times higher than the provincial average. By contrast, the average figures per capita GDP in Wuhua were much lower than the provincial average during the

1995-2005 period; the gap had increased by about 8.2 times in 2005 from 4.7 times in 1995. Although Meixian occupied the middle level of development within Guangdong, its average per capita income was also quite low in contrast to the provincial average of Guangdong.

Table 7.4: Comparison Average Per Capita GDP between Three Targeted Counties within Guangdong (unit: yuan)

Region	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
District of Guangzhou	14986	16499	18165	20054	22400	25222	27845	30685	35042	40788	46579
Meixian	3187	3626	3843	4192	4548	4880	5577	6045	6764	7521	7972
Wuhua	1118	1167	1170	1207	1279	1310	1395	1487	1597	1697	1790
Provincial average	5290	5883	6493	7139	7777	8459	9174	10099	11375	12850	14607

Note:

1. The calculation of per capita GDP was based on the constant price;
2. The figures of 2005 per capita GDP have been readjusted in order to eliminate the regional price difference within Guangdong.

Source:

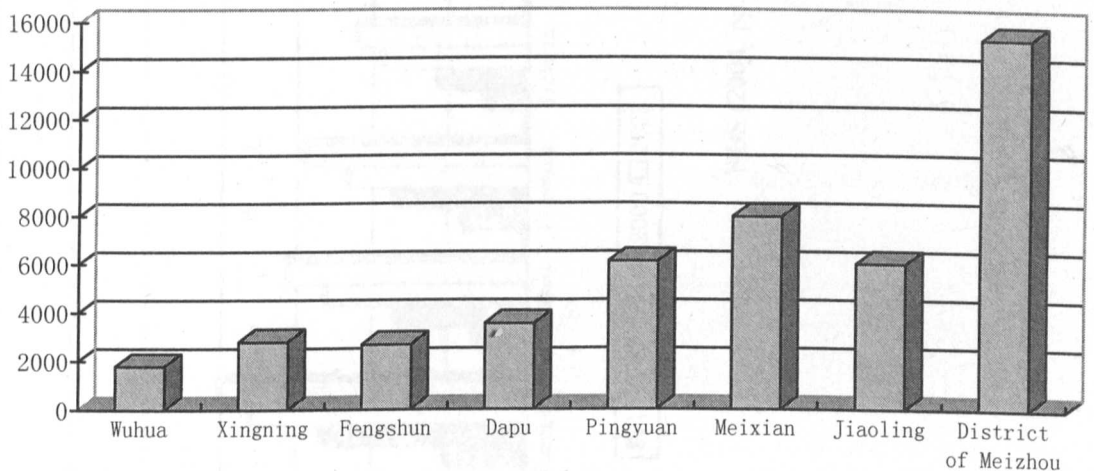
GSB (1996 – 2006, various pages)

My field research method which included interviews with officials, academic and ordinary people, was more representative, objective and in-depth than that of previous scholars. Therefore, in terms of the issue of regional income inequality in China, I believe that the findings summarized from this study will be more reliable. Wang and Hu (1999) have conducted two surveys assessing the serious degree of spatial economic inequality in China. However, only leaders from provincial, prefectural and county levels have participated in their interview surveys; no academics or local ordinary people took part in their regional disparity study. Their sample size is relatively big: as 144 governmental officers were included; nevertheless, their interviewee sample is biased due to lack of other independent voices. Moreover, at provincial level, Wei's research of regional economic development in Jiangsu (2000) was mainly based on official statistics (e.g. The Provincial Statistics Bureau of Jiangsu). Very few informal interviews were carried out by Wei; however, these interviews were used to interpret his secondary data analysis rather than further explore the personal views of these interviewees regarding local economic disparities. More generally, although both officials and researchers were involved in his interviews, opinions given by ordinary people are absent from Wei's research, as local residents of Jiangsu were not involved. Of course, opinions given by government officials are essential and useful; they can provide much detailed information with regard to state policies. Nevertheless, personal views expressed by ordinary people are also important and cannot be ignored. These local residents might have unique personal feelings about the places where they live: which do not necessarily always coincide with those given by officers and scholars. My personal judgment is that their opinions might more accurately reflect the truth about regional economies.

As far as the specific counties within Meizhou Municipality are concerned, the figures presented in Fig 7.1 show that Meixian County had a relatively high average per capita GDP. It suggests that Meixian is a relatively developed region in Meizhou; local people's average living standard is generally higher than other

regions in this municipality. In addition, in terms of average per capita GDP in 2004, Meixian County was ranked in one of the middle-level economic development positions among the counties in Guangdong at 7521 *yuan*. However, Wuhua County was not only ranked last within Meizhou but also in Guangdong as a whole in terms of per capita GDP in 2004; its per capita GDP was only about 1697 *yuan*³⁵. The average growth of per capita GDP in Wuhua was only 7.4 percent during the period between 1991 and 2005, which was much lower than in the majority of counties within Guangdong. These figures show that Wuhua was probably one of the poorest and least developed counties within Guangdong during the reform period. According to the figures published in the China Yearbook 2004 (2004, p. 935-982), the per capita GDP in Wuhua in 2005 was even lower than the per capita GDP of Guizhou and Ningxia in 2003, which were 3600 and 4997 *yuan* respectively. Guizhou and Ningxia are two of the poorest provinces in China.

Fig 7.1: 2005 Rank of Per Capita GDP in Regions within Meizhou (unit: *yuan*)



GSB (2005, various pages)

In addition, according to the data illustrated in Table 7.5, in 2000, the average per capita net rural income in Guangdong which was 3654 *yuan*, was much higher than Wuhua's per capita GDP in 2005. Moreover, in 1990, the per capita rural income in Shanghai which was 1907 *yuan*, was also higher than Wuhua's per capita GDP in 2005. This evidence suggests that economic growth in Wuhua County had been notably slow; this is not only true in relation to the other places within Guangdong but also to the other regions in China. In economic terms, Wuhua appears to be one of the poorest counties in both Guangdong and China.

³⁵ The calculation of per capita GDP in the individual regions has been based on the constant prices. Therefore, it is not necessarily consistent with the figures published in the provincial statistical yearbook.

Table 7.5: Average Per Capita net Rural Income in Selected Provinces or Municipalities in China 1990, 1995, 2000, and 2003, (Unit: yuan)



NBS (2004, pp. 291)

I believe that this sample size will not only help me to measure the degree of spatial economic inequalities, but also ensure critical analysis of the main factors causing regional inequality. Saunders, Lewis and Thornhill's study (2003) has also suggested that selecting small samples using non-probability sampling methods, sometimes can generate a meaningful, in-depth and resource-saving research study. My relatively small size was also due to limitations of time and financial resources. My field interviews in Guangdong were both a time and money consuming business; involving travel, accommodation and other social expenditure.

Chinese people do not like talking with unfamiliar people whom they regard as strangers and outsiders; let alone to engage in an interview with them. In his book *'Beyond the Chinese Face'*, Bond (1991) suggests that Chinese people are unwilling to engage in conversation with strangers, and provides research with regard to communication, habits and culture among Chinese people.

Yan's cultural research on a Chinese village named Xiajia (1996) suggests that the Chinese people treat personal relationships as important in society. Villagers enjoy communicating with familiar people, but not with strangers. Strangers, who lack personal connection with local villagers, face invisible obstacles to communication: called "socially disadvantaged positions" by Yan (1996, p. 8). Yan's argument reveals that the difficulty of conducting face-to-face interview in China should not be overlooked. In Sim's study of adolescent drug abusers and family therapy treatment in Hong Kong, China (2005), which adopted specific interview studies, he suggests that the Chinese people are not used to expressing their personal feelings and views in front of their parents; more significantly, Sim recognizes that Chinese people do not feel comfortable talking with strangers. The social workers in Hong Kong face many difficulties assessing adolescent drug abusers and prescribing medical treatment. Similarly, Sim's study demonstrates the difficulty of conducting interviews in China. Hong Kong is a classic Cantonese society which can fairly represent the general situation in Guangdong and China. It is reasonable to image that such difficulties also exist in Guangdong. This inherited communication habit is especially true of Guangdong. Ye (2005) provides a clear description of the Cantonese people. He states that,

In order to strengthen the mutual understanding, the northern people of China like going to the neighbor houses directly, and making friends with their neighbors. However, this is not the case for the Cantonese, they don't like making unnecessary social communication and talking with other neighbors ... (2005, p. 212)

Recruiting the targeted interviewees by the probability sampling method is of course an ideal approach to achieving full sample objectivity; however, Guangdong, like many other Chinese regions is a relationship-oriented society, and networks of friends and family form the basis of these relationships. Therefore, recruiting the sample interviewees directly from the targeted local population, by adopting the probability sampling method, would inevitably result in low response rates and other difficulties in accessibility. For me, the three targeted counties represented a new investigative experience. As I was unfamiliar with the local environment, it was very difficult to find, identify and access potential interviewees without

introductions from local intermediary people. This study was unable to use the probability sampling method to recruit potential interviewees. Instead, in the process of selecting local interviewees for field research, I have adopted the snowball sampling method³⁶. In my opinion, considering the Chinese and Cantonese communication culture (on which I will provide more detailed discussion in the following section) and my personal experience, snowball sampling is probably the most effective and reliable sampling method for recruiting potential sample interviewees. Furthermore, in order to obtain more reliable opinions from the targeted interviewees during the field interview, and reduce their suspicion and fears, I decided to recruit sample participants through the relationship networks of friends and family. Even if it is not a fully objective and professional sampling method, by adopting the snowball sampling, I obtained a much more successful response rate. I contacted a few potential interviewees first, and then relying on them, I identified and got in touch with other local residents. I have overcome the challenges of finding, targeting and accessing local residents through introductions from third parties, which enabled me to successfully conduct a proper and reliable interview conversation with these local people.

However, it is important to note that snowball sampling suffers from problems of sample representativeness. Bryman (2004, p. 102) argues that, “the problem with snowball sampling is that it is very unlikely that the sample will be representative of the population ...” From my viewpoint, in order to successfully access and interview the right people, obviously, the snowball sampling is the most useful method available. Saunders, Lewis and Thornhill (2003, p. 176) point out that,

However, for populations that are difficult to identify snowball sampling may provide the only possibility.

7.2.3 Research Reliability and Validity

Brink’s study (1991) suggests some essential criteria for measuring the reliability of research studies. Many other scholars have also researched this issue (e.g. Gorden, 1987; Foddy, 1993; Arksey and Knight, 1999; Bryman, 2004). Foddy is one of the most representative scholars. In order to collect reliable data during the interview, Foddy (1993) argues that it is crucial either to provide simple questions or to explain the meaning of complex research questions to the interviewees, and give them assistance in understanding the research topics; this is especially true when the interviewees face complicated words and long sentences in the questions.

In seeking more reliable and accurate interview data from the selected interviewees, I carefully considered the interviewing questions and specific words being used; any difficult and complex words were to be eliminated or clearly explained. As

³⁶ Bryman (2004) has offered a discussion on the issue of snowball sampling method. Bryman (2004, p. 100) defines snowball sampling as, “With this approach to sampling, the researcher makes initial contact with a small group of people who are relevant to the research topic and then uses these to establish contacts with others.” Similarly, when facing difficulty in identifying the targeted interviewees, Saunders, Lewis and Thornhill (2003) have also pointed out that snowball sampling is a popular research method.

mentioned earlier, due to the low level of education or illiteracy of some local interviewees, in order to ensure the research questions were understood in the same way by all participants, and that data collected in the interview could be meaningfully compared and analyzed, I clearly read all the questions and explained complex words in all field interview sections with the ordinary participants. Moreover, before the formal field interviews were conducted, pilot interviews with five targeted local residents were carried out to test the suitability and accuracy of the interview questions in August 2006. As Arksey and Knight (1999, p. 95-96) emphasize,

The next stage is to pilot the questions in an initial study; this is important even if time is short, because it is only when they are used in a trial run that you become aware of any shortcomings.

During the pilot study, any limitations or problems within the questions, discovered through feedback from the piloting interviewees, were addressed. In fact, the pilot study provided me with an opportunity to finally check the proposed interview questions. For example, the concept of 'spatial inequality' is ambiguous and the word 'spatial' was understood differently by piloting participants. Therefore, I replaced the phrase 'spatial inequality' with 'regional inequality' in all relevant interview questions.

During the three month field research experience in Guangdong, I felt that persuading the participating interviewees to freely express their real opinions and feelings during the interview conversation was one of the biggest challenges for me. As de Laine (2000, p. 2) suggests, "fieldwork becomes especially problematic when researchers cross boundaries of conventional and sensitive topics ..." China is not as democratic as the West, and it has inherited many customs and unique habits; the Confucian culture and ways of communicating with other people have been embedded in Chinese culture for more than 2,000 years³⁷. It is important to note the following point. Over the past two decades, China has experienced a dramatic change contributed by the 'reform and open-door' policies. The Chinese people have become less ideology-oriented, more pro-market and open-minded; the Chinese society is getting more open. Nevertheless, the significance of Confucius culture to the lives of ordinary citizens has remained largely unchanged. In fact, it seems to be the case that it is becoming more influential, as exemplified by popular TV programmes on Confucian culture and rapid expansion of the Confucius Institute around the world; supported by the Chinese government. In addition, the

³⁷ The common Chinese characteristics like collectivism, authority and harmony originally derived from the Five Relationships and Five Virtues of the Confucius Culture, and Confucian Work Ethic (e.g. Fan, 2000; Rarick, 2007). For example, Rarick (2007, p. 23) points out that, "The Five Relationships dictate appropriate behaviour and roles for organizational members; the Five Virtues provide a moral framework for society and stress the importance of harmony; and the Confucian Work Ethic stresses the importance of hard work, loyalty and dedication, frugality, and a love of learning." The Five Relationships include loyalty between King and Subject; Relationship between Father and Son; Duty between Husband and Wife; Obedience to Elders; and Mutual Trust between Friends (Rarick, 2007). In terms of the Five Virtues, he points out that, "In addition to maintaining harmony through relationships, Confucianism promotes five virtues: *ren*, or benevolence; *yi*, or righteousness; *li*, propriety; *zhi*, or wisdom; and *xin*, or trustworthiness." (2007, p. 25)

traditional education system in China highly emphasizes obedience and loyalty to authority: not criticism and challenge. Moreover, this education system encourages silence, cohesion and harmony rather than free speech and critical discussion. Traditionally, Chinese society has little tolerance of freedom of speech or inconsistency with accepted common social values. Western scholars have done studies on this issue (e.g. Seligman, 1990; Bond, 1991; Blackman, 1997). For example, Seligman (1990) has provided important analyses regarding the traditional Chinese education system and learning culture. Seligman argues that children are expected not to question or criticize authority, but to keep silent and accept it in order to build a homogeneous and peaceful society. In addition, owing to the general public and collective group interest, the individual person might need to give up their own opinions and follow the popular and desired social values. Seligman (1990, p. 36) highlights that,

Chinese people must listen to those in authority and do as they say. And their actions, for good or ill, reflect not only on themselves but also on all of their compatriots.

Moreover, Blackman's study (1997) offers clear comparison of common social values between the West and China. The common values of Western countries pay special attention to freedom of speech, critical discussion, human rights and personal achievements; meanwhile, the values of social harmony, collective cohesion and interests are emphasized in China. Similarly, Bond (1991) links this sort of traditional education system with long-term influence of the Confucius culture in China.

As far as Guangdong is concerned, without question, the influence of the traditional Chinese education system in this region is notable. Some scholars argue that many Cantonese are perhaps not very good at open discussion and lack relevant skills. Why are the Cantonese people not very good at expressing their own opinions? The research done by Ye has provided some explanation. To my knowledge, Ye's study (2005) is one of the few studies discussing the personality characteristics of the Cantonese people. He argues that,

The silence characteristic of the Cantonese has been cultivated during this period of long-term discrimination and oppression from the north of China. (2005, p. 48)

In my view, Ye's argument is not justified. Ye's argument might be too extreme and be strongly biased toward people in the north of China. In seeking public attention, he seems to inappropriately highlight discrimination towards the Cantonese by Northerners, without proper supporting evidence. His study seems to be just anecdotal. His argument that the Cantonese have not been taken into account by the Northerners in China is simply not appropriate. For example, Sun Zhongshan, the first President of Republic China, is a native Cantonese, even though he studied and worked abroad for many years. Ye Jianying, a key founder and distinguished leader of the People's Republic of China and Chinese Liberation Army, was also Cantonese. Therefore, Ye's study needs to be treated with caution. However, his research deserves to be considered. It highlights the silence, characteristic of many

Cantonese, and under-developed skills of argument and public discussion. Moreover, his study suggests that difficulty in persuading participants to frankly express their own perception and opinions might occur in an interview conversation. Faced with such difficulties and challenges, I have taken the following actions to try to resolve this problem, and obtain more reliable opinions and accurate interview data from the participating interviewees.

In order to strengthen the confidence and reduce the fears of participating interviewees, both I and my trusted intermediaries promised not to divulge personal information, such as the name of individual interviewees, interview locations and other sensitive personal information in the published papers and to maintain confidentiality at all times. In order to obtain the trust of the local residents, I accessed and made contact with some influential people in the local areas (such as the chef of a village). Through the introductions and promises from trusted third parties, I realized that the local interviewees were substantially less suspicious of my motives (I presume their initial reaction was probably due to my non-local background and Western education).

For example, during my field trip in Nangang, Wuhua, guided by basic research and advice from other people, Deng was the first targeted person I wanted to speak to. However, when I visited Deng's house for the first time, he did not look friendly and was very reluctant to talk with me. Although I kept talking to him of my personal background and study purpose, in his eyes, I saw doubt and suspicion. He asked me the same questions several times, such as why I came here, what I wanted, why I was interested in him. I do understand his suspicion. These least developed regions, have been ignored for too long. People like Deng, who belong to the lowest social class, have not been listened to for some years. How can they suddenly trust a stranger like me and engage in an interview conversation? Obviously, my first visit failed. I realized that establishing a personal relationship was essential to reducing Deng's suspicion and building mutual trust between us. The chef of Nangang, Wang was the key person helping me to conduct free interviews with Deng. After Wang explained to him that I was a student studying regional economy in Guangdong and conducting field research, Deng overcame his suspicion and became more open. It is clear that Deng trusts his chef. He understands that I really came to listen to his opinion regarding the regional economic gaps in Guangdong and his hardships. Several talks later, slowly, he felt comfortable enough to express himself. During the interview conversation, I was actually moved by his sincerity and natural feeling. Deng treated me as one of his friends and told me everything about him without reservation. He knew that I might not be able to help him, but at least I provided an opportunity of making his voice heard by others.

Therefore, my personal experience suggests that introductions from intermediary people play an important role in eliminating the nerves and fears of the participants taking part in interview conversations. Without personal relationships (*guanxi*)³⁸ and introductions from trusted intermediaries, it is very difficult to successfully complete field interviews. As Seligman (1990, p. 45) points out, "It's often the case

³⁸ There is a big difference between the western countries and China in viewing the importance of personal relationships (*guanxi*). In the book '*Gifts Favors & Banquets*', Mayfair Yang (1994) has provided some detailed analysis with regard to the communication culture in China and the key role played by the personal relationships in the Chinese society.

that you can't get even the simplest things accomplished in China without guanxi, but you can do just about anything – even things you probably ought not to do – when you have it.” Furthermore, Bond’s study (1991) highlights the important role played by the intermediary person.

Moreover, in my opinion, without introductions from trusted third parties, the inherited Chinese communication culture would make in-depth exploration of opinions and perceptions very difficult. Even though the people had agreed to participate in the interview, they might still exercise self-censorship and not freely or fully express their real feelings. To my knowledge, conducting field interview research is a tough and challenging job, and must be even more difficult for the foreign researcher. Of course, it is difficult for the Chinese researchers too, but without question, in terms of field study in China, Chinese native background is an advantage. Seligman (1990) rightly argues that doing business with the Chinese is made easier by the introduction of an intermediary person. Secondly, the strong distinction between the inside people (*quan nei ren*) and outside people (*quan wai ren*) has long existed in traditional Chinese communication culture. Chinese people have treated friends (as belong to the inside-group) very differently from the strangers (regarded as the outside-group); this is because the Chinese believe that friends can be trusted while strangers might betray them at some time. Blackman (1997) offers some more analysis regarding the distinction between the inside people and outsiders in the Chinese society. Similarly, Yang (1994) suggests that the aspect of the insider and outsider is based on kinship in traditional Chinese culture.

In order to build mutual trust and friendship with local residents, I have tried to act as one of them by dressing similarly to the local people, practicing local dialect and staying overnight in the different villages and townships of the targeted regions during field research. Moreover, I believe that it is crucial to treat local interview participants as friends. I should try to adopt the local common ways and just act as one of the local community. Showing my sincere care to the local participants and offering positive response to their opinions were also useful ways to encourage frank argument expression by the participating interviewees. Oakley (1990, p. 33) suggests that,

If the interviewee doesn't believe he/she is being kindly and sympathetically treated by the interviewer, then he/she will not consent to be studied and will not come up with the desired information.

I found that the elite participants liked the interviewer to listen to them and share their concerns and opinions during the interview conversation. My field experience told me that concentrating on their speech and recognizing their special positions and value were important in encouraging them to provide information and frankly offer their real opinions; they needed to know that I respected them. For example, during the interview conversation with the interviewee from Guangdong Economic and Trade Committee, by sharing his opinion and praising his achievement regarding regional economic development, I found that he was more willing to talk and provide information. Chinese officials like praise and encouragement and dislike criticism. Gorden’s study (1987) suggests that acknowledging the elites’

unique value can make the interview easier to conduct mainly due to the ego of the individual.

Moreover, besides the measures discussed above, in order to reduce the suspicion and strengthen the confidence of the targeted participants, I also provided a formal covering letter and personal statement to show my student identity and research objectives and aims, and explain to them that my research purposes were purely to analyze the regional economic disparities within Guangdong and its main causes, and bridge the research gap in this area. I found that this was also useful for building trust with participants and encouraging them to express their opinions more frankly. I have taken other actions and made certain promises to ensure the quality and accuracy of the data collected from both the local and elite interviews, and to reduce self-censoring of opinions expressed by the interviewees. For example, I used the first few minutes to have an introductory conversation with the selected local participants in order to strengthen their confidence and to make them feel comfortable. I found that it was helpful; it stimulated the more free expression of their own feelings and opinions by the interviewees, this is especially true of the main causes of slow economic growth in the least developed regions within Guangdong. Arksey and Knight (1999, p. 101) state that,

From the time of the initial meeting, which best takes place in comfortable and familiar surroundings, the researcher must attempt to establish rapport.

In addition, I have used some other techniques to achieve high response rates and cooperation from the targeted local residents. For example, reducing travel costs and time for the participants, providing small tangible rewards, such as travel expenses, free meals and gifts, all encourage the participation of potential interviewees and also encourages them to express their own opinions.

In general, despite possible shortcomings in my chosen method of recruiting the sample interviewees, I remain confident that the field interviews have provided me with a good opportunity to examine the regional economic disparities within Guangdong during the reform period. I am also convinced of the sample's general representativeness and the reliability and accuracy of data collected from the interviews. My confidence is based on four main factors. First, different levels of governmental officials have participated in the interviews: ranging from county to province. Secondly, the participating official interviewees represent various provincial government organizations. Thirdly, in order to eliminate potential bias in the opinions of the government officers, besides the interviews with the officials and scholars, I have done numerous field interviews with ordinary local people representing; feelings and opinions expressed by both the elite and ordinary people have been obtained and examined. Last but not least, the participating interviewees were selected from different regions in terms of development level.

Moreover, I have some unique advantages. Relative to the foreign scholars, I was in a much easier position conducting field research. Personal connections further ensured the smooth running of my field research. My native Cantonese background enabled me to have interview conversations with the participants in the local dialects; personal networks and relationships made it much easier to overcome the

difficulties in approaching targeted local residents and conducting field interviews. In addition, Guangdong has been one of the earliest provinces to implement the 'reform and open-door' policies, and has been opening up to academic research and foreign investors since 1978. Today there are thousands of domestic and foreign academic scholars either working or studying in Guangdong. Also, owing to influence from academic institutions in Hong Kong, the Cantonese people are more familiar with the procedures of academic research and more willing to take part in interviews than in many other Chinese regions. Although there still are various restrictions on academic research, to be fair, it is much easier to conduct field research and collect local data in Guangdong than during the pre-reform period.

7.2.4 Ethical Issues

Soltis's study (1990) suggests the importance of ethical considerations like privacy and confidentiality in protecting the people participating in any academic research. In order to analyze the main causes of regional development inequality within Guangdong, the interview participants were persuaded to give their real opinions. I believe that the true, personal and in-depth opinions expressed by the selected interviewees have played a very important role in analyzing spatial economic disparities. In order to obtain reliable and true opinions from the participants, some essential ethical measures were taken to protect the participating interviewees. Confidentiality is the soul of ethical consideration during any academic research. As de Vaus (2001, p. 87) stresses,

It is essential therefore that information be collected in such a way that confidentiality can be guaranteed.

The ethical application of my research project was approved by the Ethics Review Committee of School of East Asian Studies, University of Sheffield, in June 2006. In this study, all participating interviewees have been given full guarantees that their personal information will not be disclosed for publication and any details of third parties or people, and data and content from the interview conversation would only be used for academic research purposes.

In addition, it is also important to provide full information to interviewees regarding the aims of a research project, and the rights and consequences of their participation (Piper and Simons, 2005). Before participating in the field interviews, all selected interviewees have been given time to carefully read or be told about the 'Participant Consent Form' and 'Participant Information Sheet'³⁹; the interviewees fully understood their participation was voluntary and they could withdraw at any time without any negative consequence or punishment. The majority of the Cantonese

³⁹ Both the 'Participant Information Sheet' and 'Participant Consent Form' are a part of University Research Ethics Application Form. According to the guidance given by the Research Ethics Committee of the University of Sheffield, 'Participant Information Sheet' should give the full information regarding the research project like the purposes and aims of research; therefore, the potential participants can make a proper decision about their participation in the interview. The statement of free withdrawal from interview participation during any research period should also be clearly identified in this information sheet. Moreover, information in the 'Participant Information Sheet' should include the place of interview, the research purpose and report of research findings.

interviewees did not want to sign the 'Participant Consent Form' for the confidentiality and safety reasons; however, all of the participants have given their verbal consent to take part in the field interview conversation. In order to build trust with the participating interviewees, encourage them to express their real opinions, and reduce self-censorship, I only took notes when necessary and concentrated on listening to their opinions without using any voice recorder machines during the interview conversation.

The interviewees' personal views may inconsistent with or even challenge the official position and arguments made by the government. Considering the relatively sensitive issues and negative topic of regional economic disparity in China, the past political and social campaigns like Cultural Revolution, contemporary authoritarian and less-transparent political system, and the powerful and unaccountable governmental officers⁴⁰, such disagreement might cause potential political and economic disadvantages or negative consequences to the interviewing participants. In addition, as discussed earlier, there is lack of active NGOs⁴¹ and other independent media in contemporary Chinese society; thus, the actions taken to protect the rights of the participating interviewees and protect them from harm are more prominent. As Yang (1994, p. 294) suggests,

Given the difficulties of independent group formation in the present Chinese political context, the inability of the legal system to protect individuals and associations and the relative weakness of a discourse of "rights" and of "individualism" in the culture, the prognosis of civil society according to the European model does not look good in China.

Therefore, in order to fully protect the rights of interviewees and eliminate any potential harm to them, the personal information of interviews would be kept confidential and secret all the times; the names and addresses of individual interviewees and interview locations have been purposely removed or replaced with pseudonyms. I hope that the reader will be aware and understand that these necessary precautions were taken for the protection of participating interviewees.

7.3 Conclusion

Secondary analysis of previous literature and statistical data is unable capture the full picture of the research topic, and cannot discover all the important state-oriented regional development policies during the reform period. More significantly, this secondary analysis cannot reflect personal perceptions and

⁴⁰ Cai's research (2000) has provided more detailed analysis regarding the issue of lack of accountability for the Chinese officials during the less-transparent promotion system in the government.

⁴¹ In terms of the NGOs' development in China, He's study (2006) suggests that the Chinese NGOs have been entirely controlled and supervised by the local governments and the NGOs have not yet played the important role of the independent body in China and actively participated in local political, social and economic affairs. Similarly, the study conducted by the UNDP (2000, p. 43) clearly suggests that, "However, there are some unique characteristics to civil society in China during the transition. The most important one is that most of the new institutions are not fully independent at all, but rather are semi-official organizations."

opinions given by the participants without field interview. The field research interviews have given me personal perception of the seriousness of regional economic disparities within Guangdong, and greater opportunities to investigate and measure the issue of regional inequality and its main causes. In fact, the fieldwork research has given me a meaningful and rewarding experience. Due to the limitations of questionnaires, I have adopted the interview approach to collection of field research data. Moreover, snowball sampling has been used to recruit potential interviewee; as the guanxi-oriented communication custom in China makes direct probability sampling extremely difficult to conduct. In order to encourage the interviewees to express their opinions freely, and ensure the quality and reliability of data collected from the field interviews, I have adopted various measures, techniques and actions. In terms of the ethical issues of this study, confidentiality is the key. In order to protect the rights of the interview participants and prevent them being harmed in any way all private information disclosed in the interviews would remain highly confidential.

Chapter 8 Chinese Data, and Associated Limitations

8.1 Introduction

This chapter attempts to provide a critical discussion of the literature review on the quality and accuracy of Chinese statistics, particularly those concerning GDP per capita and its annual growth rate. Section Two examines the reliability of Chinese statistics; I offer an in-depth discussion on the quality of Chinese statistical data and the capability of official departments to handle statistical data. Moreover, the specific issues of statistical quality in Guangdong is also addressed in this section. The final conclusion are drawn in Section Three. The key questions I am keen to investigate are the following: how reliable is the Chinese statistical data? Is it safe to utilize Chinese statistical figures in academic research? Is it justified to argue that the official data is completely useless?

8.2 Analysis of Official Statistical Data in China

The study conducted by Arksey and Knight (1999) is important; they stress the importance of official statistics and other written materials to academic research. In terms of this research project, I have accessed and utilized comprehensive statistical data and documents published by both the national and provincial government in China. In addition, selected specific local documents and historical records (*di fang zhi*)⁴² from Guangdong have also been widely cited. Many other written materials and documents released by the authorities have become accessible for analyzing the spatial development disparities within Guangdong. The provision of relatively rich regional statistical data and other government materials has greatly encouraged me to examine spatial economic inequality within Guangdong during the ‘reform and open-door’ period.

Moreover, due to its comparative ease of collection, this sort of secondary analysis of official data has saved my limited time and financial resources. The analysis of official statistical data has allowed me to concentrate on measuring the regional economic inequalities within Guangdong, and examine the main causes of this phenomenon. The main questions I address in this section are these: how reliable and accurate is the Chinese statistical data? Can we trust these official figures? More specifically, relative to other Chinese provinces, what is the quality of statistical data in Guangdong?

8.2.1 The Quality of Chinese GDP Statistical Data: Its Critics

Official statistical data has been widely criticized in recent years. Many scholars have raised doubts and suspicion about the reliability, accuracy and quality of Chinese official statistics. For example, in the early study, Orleans (1974) argued that the reform and improvement of Chinese statistical system was unsatisfactory and too slow. In addition, he implicitly indicated that Chinese statistical officers might deliberate falsify statistical data in order to receive personal benefits.

⁴² “Local gazetteers were already a distinct genre of history writing during the Song dynasty (960-1279), and China has a continuous tradition for compiling information on all aspects of local natural and human conditions in this form.” (Thøgersen, 2006, p. 191)

Similarly, Cai Yongshun's study also strongly doubted the credibility and honesty of Chinese statistical officials and particularly of the masses of rural statistical personnel, and blamed China's political and statistical reporting system. Cai (2000) firmly believes that statistical falsification and manipulation is inevitable in China; he assumes that local statistical officers have purposely manipulated and embellished data in order to pursue their own political and economic interests. As Cai (2000, p. 803) argues,

As long as local officials have a stake in statistical reporting, the manipulation cannot be avoided if effective checks are not instituted ...

For many years, national statistics data of China have heavily relied on the raw data collected and reported by local authorities. In fact, this hierarchical reporting system has brought many opportunities for the local government officers to falsify and embellish statistics data in order to pursue personal interests. Rawski (2000, p. 6) believes that "the traditional system of vertical reporting meant that NBS officials had neither the authority nor the capacity to correct errors or distortions arising at lower levels within the reporting system." Moreover, Xue's study (2004) suggests that the hierarchical and complex statistical reporting system in China has made easier the exaggeration and manipulation of data by performance-motivated local government officers. Rawski is probably one of the most critics regarding the quality of Chinese statistics data. Rawski (2000) strongly criticizes the quality of Chinese statistical data. He argues that the limitations and incapability of the statistics bureaus of China inhibit improvement of the statistical system, and make it unable to produce accurate and reliable economic statistical data. In particular, Rawski (2000) believes that the issue of Chinese statistics exaggeration has become particularly severe since 1998. Rawski argues that the Chinese GDP growth rate should be 5.7 percent or so rather than the 7.8 percent claimed by the National Statistics Bureau of China (NBS) in 1998. Moreover, Rawski (2000) indicates that the real Chinese GDP growth rate might be even lower than 5.7 percent due to incomplete and missing statistics relating to wages and profits in rural enterprises. In a later study, Rawski (2001) further criticizes the overestimation, falsification and inconsistencies of Chinese official data. He predicts that the accumulated growth rate of GDP between 1997 and 2001 should be reduced to around one-third percentage or even less of the equivalent figures announced by the NBS.

On the one hand, many criticisms toward the quality of China's GDP statistics are concentrated on Chinese industrial growth. For example, Wang and Meng (2001) emphasize their criticism of Chinese GDP on industrial growth. Wang and Meng argue that the NBS has overestimated the real annual industrial growth rate by about four percent during the 1990s. Their study suggests that a more accurate figure would be 7.9 percent between 1978 and 1997 instead of around 9.8 percent claimed by the NBS. Moreover, Harry Wu (2002), through his own systematic empirical test, stresses that the official figures have significantly overestimated the actual industrial growth.

On the other hand, declining energy consumption in China during the reform period has caused some scholars to doubt the reliability of Chinese GDP statistics. In the light of fast declining energy utilization, many scholars argue that the real growth of Chinese GDP has been significantly overestimated. For example, Meng and Wang's

study (2001, cited in Holz, 2003) argues that the growth rates of industry and total gross domestic products (GDP) have been overstated since 1992 based on the comparison between energy utilization, the output of commercial production, and the growth rates of industry. Moreover, Sinton's study (2001) has demonstrated that the quality and accuracy of Chinese energy data has decreased since 1995.

The arguments of these critics demonstrate that actual GDP statistical figures have been significantly manipulated by the officials due to the decline of energy consumption in China. Therefore, the key question is whether or not a coexistence of both dramatic increase of China's GDP and fast decline of energy consumption is possible. The answer to this question is probably 'Yes'. The dramatic increase of international energy prices in recent years, the technological innovation and ownership reform of the Chinese state-owned industrial enterprises might be the three key forces contributing to the fall of energy consumption in China; the state-oriented policies of efficient energy management and energy saving have probably also played an important role. In fact, in their early study, Sinton and Fridley (2000) recognize that structural reform of industry and state-oriented energy reduction policies have attributed to the fall of energy use in China. Moreover, Fisher-Vanden, et al, (2003, p. 1) point out that,

Changing energy prices, research and development expenditures, and changes in ownership and industry composition are found to be the principal drivers of reported declines in energy intensity and consumption in Chinese industry.

Jefferson, et al, (2002) found the relationship between technological innovation and industrial productivity to be close; their statistical model demonstrates that the decline of energy consumption of industrial firms have dramatically reduced their energy consumption due to technological innovation and increase of R&D investment. In addition, hidden and illegal energy consumption by Chinese industrial enterprises, such as oil smuggling and purchasing of coal from small and illegal private coal mines, which were officially forced to shut down, have clearly not been calculated in official data of energy imports; but they have been included in the energy consumption statistics published by the statistical bureau. Under these circumstances, the actual energy consumption should be higher than the equivalent official figures. Sinton and Fridley's research (2000) has offered a detailed discussion of this issue.

Another severe critic of Chinese GDP statistics is Maddison. Due to the abolishment of statistical departments and dismissal of skilled statistical personnel, Maddison (1998; 2006) argues that past political campaigns such as Cultural Revolution caused huge damage to the Chinese statistical reporting system in the pre-reform period. In his early study, Maddison (1998) offers detailed confrontation measurement of economic growth rates between 1952 and 1995. In terms of the period 1978-1995, the annual growth rate of GDP per capita given by official statistics department was 8.4 percent. By contrast, the equivalent growth percentage of 6.0, presented in his measurement, was a striking 40 percent lower. In contrast to official growth figures, his main measurement disparities given are in the industrial and service sectors rather than in agriculture. For example, the annual growth rate of value added in the industry measured by Maddison was only 8.5 percent between

1978 and 1995; however, the equivalent figure released from the officials was up to 12 percent. As far as the statistical data of tertiary services is concerned, the annual average growth rate of value added in 'non-productive' service presented in the official measurement reached 11.8 percent; meanwhile, the equivalent figure shown in Maddison's own measurement was only 6.7 percent. (Maddison, 1998, p. 160)

Moreover, Maddison (2006) points out that both this inherited poor statistical reporting system and long-term influence of the Soviet material production systems (MPS) on Chinese statistics practice have made it difficult for China to catch up with international standards of statistical reporting during the 'reform and open-door' periods. Maddison (1998) highlights that the key differences between his and the official statistical measurement approaches lie in sectoral measurement techniques and weights. He argues that his approach to Chinese GDP measurement is more objective and accurate. Maddison (2006) seems to argue that the official statistics cannot possibly be accurate, according to his 1990 estimation. He demonstrates that the official GDP statistics were unbelievably low in 1952 and too high in 2003. However, it is important to note that Maddison's re-estimation of the Chinese GDP statistical figures faces strong criticism from numerous scholars. Carsten Holz is one of the most representative scholars. Holz (2004) claims that Maddison's own estimation of Chinese industrial growth between 1978 and 1995 (8.56 percent compared to the official 12.02 percent) is highly dubious due to his unreliable adjustment procedures. Holz (2004) further argues that Maddison's re-estimation of Chinese GDP growth during the 1978-1995 period is inaccurate, due to under-evaluation of actual growth rate in 'the secondary and tertiary industries. Holz thinks that the justification of downward growth evaluation of the 'other services' by Maddison is unreliable because of omission of employment figures and the unjustified evaluation of zero growth of labour productivity in the 'other service' sector, as Holz referred to it.

Maddison's own estimation of Chinese industrial growth was based on Harry Wu's (2002) earlier study. Wu argues that Chinese inflation has been underestimated by the official deflator and the actual industrial growth has been exaggerated (Holz, 2005). However, Wu's estimation is not convincing because there is bias existing throughout his empirical tests. As Wu himself admits the shortcomings of his estimation approach, he says, "Some biases remain in our estimates." (2002, p. 194). In considering the product method adopted by Wu, Holz (2005) believes that none of his four arguments criticizing official industrial data are reliable and valid. Holz (2005) does not believe Wu's official deflator criticism because the NBS's published deflator regarding value-added industry was not based on the direct collection of product quantity statistics from the rural and local enterprises (NBS Industry and Communication Division, 1999, cited in Holz, 2005). In addition, Zheng Jingping, a senior statistical officer from the NBS also argues that sample surveys conducted by its own survey team of the NBS have been dramatically expanded during the reform period. Through sample surveys the NBS has directly collected statistical data in various areas, such as population, price, retail and small industrial firms and household (Zheng, 2001). Therefore, Wu's criticism toward the Chinese statistics is not very reliable or accurate. As it is based on Wu's problematic previous calculation and estimation, Holz suggests that Maddison's readjustment is hardly justified. Holz (2006) further criticizes the alternative estimation of China's GDP growth provided by Maddison, based on his four

re-adjustments of NBS statistical data, as highly problematic and not convincing. For example, the alternative employment data claimed by Maddison might underestimate the actual employment growth of the service industry and overstate the employment growth among military personnel. In addition, Maddison's own labour productivity growth in China is also unreliable⁴³ (Holz, 2006). Holz makes a conclusive judgment on Maddison's adjustment of China's industrial and GDP growth, he states that,

I have, beyond arguing/showing that the operational methods chosen by Maddison for making adjustments do not hold up to scrutiny, in my view also shown that the specific reasons advanced by Maddison (1998) to justify his adjustments are problematic. (2006, p. 473-74)

In response to Holz's criticism, Maddison (2006) emphasizes that it is necessary and also common international practice to re-evaluate official statistics in order to make comparisons with different countries and eliminate official bias. Maddison (2006) points out that Holz is far less critical and sceptical than he is regarding the official statistical data. He firmly argues that some important issues such as purchasing power parity (PPP) and the accuracy of China's GDP growth in time-series comparison have been entirely unaddressed by Holz's study.

In my view, Holz might have too much sympathy toward the work done by both national and local statistical bureaus in China. Holz acts as a strong defender of Chinese official statistics. However, it is important to note that Maddison is far more sceptical of Chinese statistical data and too pessimistic about the quality of the Chinese GDP statistics. He seems to understate the improvement and dramatic change of the Chinese statistical system over the last two decades. I have provided more discussion regarding the change of Chinese statistical practices in the following section. Of course, it is the case that the national statistical data in a country should be re-examined and adjusted in order to eliminate potential bias and make valid international comparisons; but it is inappropriate to adopt Maddison's strategy of comparing Chinese statistical data with the equivalent figures in western developed countries like the OECD countries. China's traditional statistical system is dramatically different from those of western countries, in areas such as the statistical definition of the secondary and tertiary industries; this is especially true of the statistical system before 1978. Maddison did not address these issues in his research. Therefore, his over-simple conclusion with regard to the unreliability and inaccuracy of Chinese GDP statistics is unfair and unwarranted. As Holz (2006, p. 472) suggests, "I argue that OECD countries are unlikely to be good comparison countries for China, an economy in transition and at a very different stage of development." In addition, Maddison's zero growth assumption of Chinese labour productivity in the non-productive service sector during the 1990s is odd and without reasonable explanation. The dramatic increases of Chinese industrial and total GDP obviously cannot be attributed to capital investment alone. Although the

⁴³ However, in response to Holz's criticism of his data on labour productivity growth, Maddison (2006, p. 124) argues that, "Holz considers this an egregious error, and suggests I should have assumed a 5-6 percent per annum increase in labour productivity. However, the evidence he produces does not warrant such a conclusion."

actual growth of Chinese labour productivity might be still unknown, it has experienced steady increase during the reform period.

Moreover, Maddison's claim that Holz offers a less critical examination of Chinese statistical data is unfair. Holz has, indeed, done important research on the quality of Chinese GDP statistics, and he has made strong criticism with regard to the quality of the revision data published by the NBS and the capability of the NBS in collecting and checking local raw statistical data. For example, Holz (2003) has raised criticism about the accuracy of the official data in household consumption: which occupies about 50 percent of total national GDP. He reconstructed figures of household consumption based on official GDP statistics. Significantly, he found many problems such as the inconsistency of official GDP explanations and household consumption data during the period between 1997 and 2001. Therefore, Holz concludes that it might be difficult and invalid to conduct systematic cross-year comparison of Chinese GDP data.

Before offering more discussion regarding Chinese statistics and my conclusion, it is first important to examine the revised GDP results published in the 2004 China economic census.

8.2.2 The 2004 China Economic Census

The NBS published the revised results of national total GDP in the 2004 economic census⁴⁴. Relative to the pre-revision statistical data, the 2004 China economic census conducted by the NBS has made dramatic upward revision of total GDP of China and its annual growth rate. The data presented in Table 8.1 demonstrates that total GDP had increased dramatically year by year from 1993 to 2004. In 2004, the revised total amount GDP was 2300.2 billion *yuan* more than the equivalent pre-revised figure. In 1993, the difference of total GDP between the revision and pre-revision data was only 70 billion *yuan*. In addition, according to the data shown in Table 8.3 below, the revised percentage of annual GDP growth rate reached around 9.9 or so up by 0.5 percent from 9.4 during the pre-revised period between 1993 and 2004.

In addition, two prominent aspects need to be highlighted in the revision data of GDP demonstrated in the 2004 economic census. First, in comparison of the revised total GDP, with the GDP changing trend presented in Fig 8.1, the total GDP gap between pre-revision and revision had increased significantly between 1998 and 2004. Secondly, in contrast to the pre-revision data, the output of primary and secondary industries in total GDP remains quite similar; the main data revision was concentrated on the share output of the tertiary industry in total GDP. According to the figures illustrated in Table 8.1, the statistical difference of the primary and secondary industries between pre-revision and revision was quite small between 1993 and 2004; the biggest gap was the output of secondary industry in 2004, which was only 151.7 billion *yuan*. The figures shown in Table 8.2 demonstrate more clearly that the change of growth rate in the primary and secondary industries was

⁴⁴ The 2004 economic census conducted by the NBS has systematically re-measured and re-calculated the total amount of Chinese GDP and its three components of primary, secondary and tertiary sectors during the period between 1993 and 2004.

zero between pre-revised and revised figures. However, the output of tertiary industry had experienced a significant change between 1993 and 2004 in the revised figures. The output difference of tertiary industry jumped to 2129.7 billion *yuan* in 2004 from only 66.8 billion *yuan* in 1993. The average annual growth rate of tertiary industry was about 50 percent between 1993 and 2004 in the revised data.

Table 8.1: Change of Total Amount of Chinese GDP after the 2004 Economic Census (unit: billion yuan)

Year	Revised data			Pre-revised data			Data change					
	GDP	Primary	secondary	Tertiary	GDP	Primary	Secondary	Tertiary	GDP	Primary	Secondary	Tertiary
1993	3533.4	688.7	1645.5	1199.2	3463.4	688.2	1642.8	1132.4	70	0.5	2.7	66.8
1994	4819.8	947.1	2244.6	1628.1	4675.9	945.7	2237.2	1493	143.9	1.4	7.4	135.1
1995	6079.4	1202	2868	2009.4	5847.8	1199.3	2853.8	1794.7	231.6	2.7	14.2	214.7
1996	7117.7	1388.6	3383.5	2345.6	6788.5	1384.4	3361.3	2042.8	329.2	4.2	22.2	302.8
1997	7897.3	1426.5	3754.3	2716.5	7446.3	1421.1	3722.3	2302.9	451	5.4	32	413.6
1998	8440.2	1461.8	3900.4	3078	7834.5	1455.2	3861.9	2517.4	605.7	6.6	38.5	560.6
1999	8967.7	1454.8	4103.4	3409.5	8206.7	1447.2	4055.8	2703.7	761	7.6	47.6	705.8
2000	9921.5	1471.6	4555.6	3894.3	8946.8	1462.8	4493.5	2990.5	974.7	8.8	62.1	903.8
2001	10965.5	1551.6	4951.2	4462.7	9731.5	1541.2	4875	3315.3	1234	10.4	76.2	1147.4
2002	12033.3	1623.9	5389.7	5019.7	10517.2	1611.7	5298	3607.5	1516.1	12.2	91.7	1412.2
2003	13582.3	1706.9	6243.6	5631.8	11739	1692.8	6127.4	3918.8	1843.3	14.1	116.2	1713
2004	15987.8	2095.6	7390.4	6501.8	13687.6	2076.8	7238.7	4372.1	2300.2	18.8	151.7	2129.7

(Source: NBS, 2006, available on the official website of NBS)

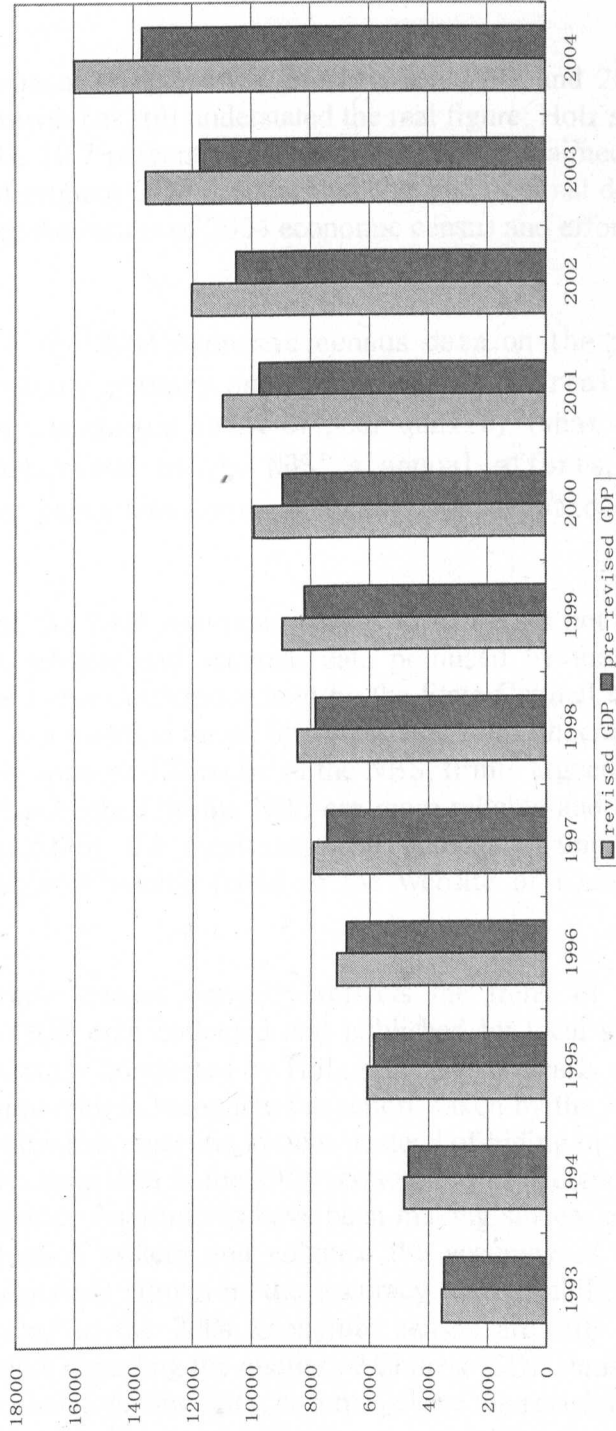
Table 8.2: Change of GDP Growth Rate in China after the 2004 Economic Census (unit: percentage)

Year	Revised data			Pre-revised data			Data change					
	GDP	Primary	Secondary	Tertiary	GDP	Primary	Secondary	Tertiary	GDP	Primary	Secondary	Tertiary
1993	14	4.7	19.9	12.1	13.5	4.7	19.9	10.7	0.5	0	0	1.4
1994	13.1	4	18.4	11	12.6	4	18.4	9.6	0.5	0	0	1.4
1995	10.9	5	13.9	9.8	10.5	5	13.9	8.4	0.4	0	0	1.4
1996	10	5.1	12.1	9.4	9.6	5.1	12.1	7.9	0.4	0	0	1.5
1997	9.3	3.5	10.5	10.7	8.8	3.5	10.5	9.1	0.5	0	0	1.6

1998	7.8	3.5	8.9	8.3	7.8	3.5	8.9	8.3	0	0	0	0
1999	7.6	2.8	8.1	9.3	7.1	2.8	8.1	7.7	0.5	0	0	1.6
2000	8.4	2.4	9.4	9.7	8	2.4	9.4	8.1	0.4	0	0	1.6
2001	8.3	2.8	8.4	10.2	7.5	2.8	8.4	8.4	0.8	0	0	1.8
2002	9.1	2.9	9.8	10.4	8.3	2.9	9.8	8.7	0.8	0	0	1.7
2003	10	2.5	12.7	9.5	9.5	2.5	12.7	7.8	0.5	0	0	1.7
2004	10.1	6.3	11.1	10	9.5	6.3	11.1	8.3	0.6	0	0	1.7

(Source: NBS, 2006, available on the official website of NBS)

Fig 8.1: Comparison of Total GDP of China between Pre-revision and Revision Figures in 2004 Economic Census (unit: billion yuan)



(Source: NBS, 2006, available on the official website of NBS)

If the revised GDP data of China announced in the 2004 economic census are more accurate, it suggests that the pre-revised data published by the NBS significantly underestimated total GDP and its actual growth rate; this is particularly relevant to the output of the tertiary industry. The unclear definition of tertiary industry and technical difficulties are probably two important reasons contributing to this underestimation of GDP by the NBS. The key question is which set of GDP data is more reliable and accurate between the pre-revision and revision figures disclosed by the 2004 economic census? Again, this question has caused more doubts and criticisms of the quality of Chinese statistics and the capability of handling data by the NBS. For example, due to the changing factor of nominal values in the various sectors of industry, Holz (2006) argues that the unchanged output of secondary industry in the revised data is not convincing or reliable.

Holz also criticizes the revised annual GDP growth rate between 1993 and 2004. He suggests that the revised annual growth has still understated the real figure; Holz seems to argue that the actual rate might be 10.7 percent rather than 9.9 percent claimed by the 2004 economic census in terms of original GDP deflator and changed nominal data. In a conclusion, Holz strongly criticizes the results of 2004 economic census and efforts made by the NBS. He claims that,

The fact that NBS did not use the 2004 economic census data on the primary and secondary sector to revise the primary and secondary sector real growth rates could be interpreted as the census being of poor quality (what a waste of money and effort!) in comparison to the NBS' s annual efforts, or of political (?) decisions taking precedence over accuracy in statistical data. (2006, p. 4-5)

In addition, the revision results of the 2004 economic census in China do not seem to support Xie's argument of more reliable and accurate data produced by the Central Statistical Bureau of China. In the Press Conference held by the State Council of China on 25th January 2007, in response to a question raised by a journalist from Phoenix TV of Hong Kong, Xie Fuzhan, the newly appointed Director of the NBS, firmly argued that the GDP statistical data collected and published by the NBS are more reliable and accurate than the equivalent data announced by the local statistical bureaus: mainly due to assistance in calculation from China Customs (cited in the website of China News Network, 25th January, 2007).

Significantly, if the 2004 economic census correctly reflects the trend of Chinese economic development, then the GDP data collected and published by local statistical department are more reliable. The study conducted by Holz (2006) also clearly suggests this point. My view is this: it is important to recognize the actions taken by the NBS and the improvement of the Chinese statistical reporting system. Instead of hiding or ignoring the problems of data quality, the revision data in the 2004 economic census demonstrates that both the central and local statistics departments have been making sincere efforts to improve the poor statistical collection system and enhance the accuracy of Chinese statistical data during the reform period. However, the accuracy and reliability of the revised statistical data demonstrated in the 2004 economic census are still open to question and have raised more doubts regarding the quality of Chinese GDP statistics and the capability of the NBS in data collection. One thing becomes clear: the revision data of

the 2004 economic census has indisputably shown that Chinese GDP statistical data did suffer from problems of quality, accuracy and reliability during the reform period; these problems have become particularly prominent since the year of 1997 or 1998. Moreover, the revision figures presented in the 2004 economic census suggests that the reliability and accuracy of GDP data in the tertiary sector is obviously far lower than in the primary and secondary industries.

8.2.3 The Quality of Chinese GDP Statistics

It is certainly the case that some Chinese GDP statistical data have suffered from problems of low quality and inaccuracy; it is undeniable and widely agreed by the academic community. On the one hand, some limited local statistical data have been purposely manipulated and embellished by the government officials; this is especially true of the rural areas. On the other hand, some figures collected by local staff have lacked accuracy and reliability, due mainly to poor statistical techniques, shortage of professional statistical personnel, and lack of sufficient fiscal investment in the statistical sector by the government. Moreover, recently, following accusations of corruption and a personal scandal, the former Director of the NBS, Qiu Xiaohua, has been arrested by the police. Considering that the Head of the NBS is involved in corruption, how can the people trust the quality of Chinese statistical data published by the NBS and even the general credibility of the NBS? In particular, the inconsistency of the Chinese GDP statistical data makes a time-series comparison and analysis of Chinese economic development difficult to conduct; As Holz and Lin (2001a, p. 29) point out,

Yet, as this paper shows, definitions of variables and the categorization of enterprises have changed frequently in the two decades of reform. ... These changes, most of which have gone unnoticed in the literature, have caused severe comparability problems in both time series and cross-sectional data.

Holz (2003) further stresses that this problem of time series data comparison is mainly due to inconsistency of the statistical indicators and definition made by the Statistics Bureau of China. In fact, many critics have concentrated their criticisms on the data inconsistency between the central and local statistics departments. The striking GDP data difference between the central and local regions since 1996 or so has caused many doubts of the reliability of Chinese statistics. How can this dramatic data discrepancy happen? How can we trust the quality of Chinese statistics if the different statistical departments themselves cannot produce a uniform set of statistical data? Holz's study (2006) highlights the argument that data inconsistency between the central and local statistics departments had become severe during the period between 1996 and 2004. The combination of provincial GDP had been more than 19 percent higher than the total GDP announced by the NBS by 2004.

However, in reply to the question asked by a journalist from Phoenix TV of Hong Kong, Xie Fuzhan (2007), Director of NBS, has offered more detailed explanation with regard to the inconsistency between central and local statistical departments. Xie firmly argues that the separate systems of data calculation of the central and local statistical departments and the double-counting output of the multinational firms are the two key reasons (Press Conference held by State Council of China on 25th January 2007, cited in the China News Net, 15th June, 2007). Xie's argument has only partly answered the

criticisms as this problem might also have been caused by other important factors, such as the actual output value of the township and village enterprises (TVEs), the profits of the state-owned enterprise sectors, and foreign trade. To blame double-counting output of multinational corporations alone for causing data inconsistency is unfair and hard to accept; from my point of view, it is an over-simplification. Moreover, the statistical department in China has frequently changed statistical definitions of sectors such as the secondary and tertiary industries and energy consumption without proper and clear explanation; this might also have caused the central and local inconsistency of statistical data. Holz and Lin (2001b) blame the less than transparent indicator definitions and confused categorization of organizations by the NBS.

Moreover, in replying to my question regarding the accuracy of Chinese GDP statistics and the data inconsistency between central and local governments, Yao Jingyuan, the chief economist of the NBS, also attempts to defend the accuracy of statistical data published by the NBS⁴⁵. He cites the example of the weather forecast to demonstrate the natural statistical difference between the central and local governments, for example, the weather forecast reported by Beijing is always different from the weather report from the national meteorological bureau of China. He argues that,

The data collection and analysis conducted by the NBS relies on its own three rural, urban and enterprise statistical survey teams (*tongji diaocha dui*). ... the statistics survey conducted by these three teams covers total 200 municipalities and 800 counties among 31 provinces in China. (Yao, 2007, cited in 'Rising China in the Age of Globalization' Inaugural International Conference)

Again, Yao's argument fails to totally alleviate the quality concerns regarding Chinese GDP statistics. It is of course the case that funding and direct control of national statistical survey teams by the NBS can ensure their independence and effectively reduce local intervention in daily statistical collection. However, in my view, the working boundaries of the national statistical surveys need to be far larger: in order to generate representative sample data which is both reliable and objective. It is important to note that there is still a significant number of Chinese regions, enterprises and economic activities that have not been covered by these three national statistical survey teams. There are more than 600 municipalities and 2100 counties within China. At most, the working boundaries of the three national teams only cover one-third of Chinese areas; in other words, the statistical reporting and collection of the other two-thirds of Chinese regions by the NBS still depends on local governments, which frequently use the traditional experiencing, sampling and prediction statistical methods in data collection. Therefore, the statistical quality of these regions is far more uncertain and may not pass strict scrutiny. Under the circumstances, the quality of statistical data published by the NBS might not be much higher than the local statistics if they are not entirely similar.

The crucial questions I am interested in are these: is Chinese GDP statistical data completely pointless? Do these numerous criticisms justify concluding that the Chinese statistical figures are entirely unreliable and inaccurate? The answer to these questions is probably 'No'. The shortcomings and limitations of its statistical system do not

⁴⁵ The accuracy question of the Chinese GDP statistics was raised by the author during the 'Rising China in the Age of Globalization' Inaugural International Conference of the UCD Confucius Institute for Ireland and the Irish Institute for Chinese Studies, held in 16th – 18th August 2007, Dublin.

necessarily mean that Chinese statistical data is completely inaccurate. It does not suggest that the scholars cannot refer to Chinese statistics in their academic research. Wang and Hu (1999) point out that it is safe to cite the Chinese data if it shows the general trend of economic development.

As discussed earlier, numerous scholars have argued the poor quality of the Chinese statistical system and practice, and the prevalent statistical falsification and embellishment by government officers. In terms of statistical manipulation, in my opinion, it is unfair and totally unjustified to argue that statistical data are entirely manipulated and embellished by the officials. Without any doubt, Chinese economic growth has been remarkable over the last two decades. Chow (2005) offers a fair judgment of Chinese statistical data by suggesting that Chinese statistics are largely credible.

It is important to note that Chinese statistical data are generally reliable and accurate. More significantly, the data reflects the basic trends and general reality of the fast growth of the Chinese economy and industries. Annual 9 percent economic growth rate claimed by the NBS or 5.7 percent or so argued by the critics, it is undeniable that either of these growth rates of Chinese GDP is impressive and relatively high in crossing country comparisons. Klein and Özmucur's empirical estimation (2002/2003) has examined some important indicators which are consistent with fundamental accounting principles to test the reliability of Chinese statistics. They found that all important components of these indicators had demonstrated the accuracy of the trend of Chinese economic development claimed by the official evaluation. Some limited inaccuracies in data would not sufficiently challenge the general legitimacy of the Chinese statistics or change the overall picture of dramatic economic growth in China. Chow (2005, p. 12) argues that,

First let it be noted that for the purpose of studying long-term trends we can tolerate sizable inaccuracies on the levels of the variables. ... Even if the overestimation is not proportional, say by 10 percent in 1978 and by 20 percent in 1998, the average rate of growth estimated by comparing the levels of GDP in these two years will not be affected substantially. This point is a matter of arithmetic.

In addition, although Chinese GDP statistics may have the limitations discussed in the previous section, some criticisms are pure speculation. Without sufficient hard evidence to support these criticisms, why should the people believe that the Chinese statistics are entirely inaccurate? Why should they believe these highly dubious arrangements rather than statistical data collected and checked by the many hard-working statistical personnel in China? More significantly, staff from the official statistical bureaus have access to more first-hand statistical information than their critics; and they have generally more accurate knowledge with regard to actual economic development in China. Moreover, many scholars seem to misunderstand and misinterpret the Chinese GDP, industrial and other data. For example, Rawski's study is not above criticism; in fact, his approach is problematic. Rawski did not clearly explain why he had chosen the income approach instead of other approaches, and he did not seem to sufficiently aware of the shortcomings of the income approach in calculating the Chinese total GDP. His income approach might have underestimated the real growth of Chinese GDP and omitted some important variables. Holz (2003) also shares this point and makes strong criticisms of Rawski's calculation of Chinese GDP growth.

In addition, Albert Keidel's study has used the expenditure approach to re-estimate the growth rate of Chinese GDP between 1995 and 2000. He claims that the official figures have overstated the actual GDP growth rate. The readjustment data of GDP growth calculated by Keidel are quite different from the official figures; Keidel's growth figures of China's GDP are lower than the official figures in some years while higher than official figures in other years (Holz, 2003). However, Keidel's expenditure approach of GDP estimation is not convincing. In fact, it is highly unreliable. Keidel's study might have undervalued the actual consumption expenditure of China, due, for example, to omission of expenditure of some household consumption and lack of calculation of consumption expenditure in some private and rural sectors. In addition, owing to the backward local statistical reporting system, many people recognize that it is very difficult to collect all of the economic consumption activities in the Chinese rural areas; this might be particularly serious in the underdeveloped west and southwest regions. Holz (2003, p. 152) also points out that,

But the expenditure approach GDP data are not without serious problems, and therefore are highly unlikely to constitute the "correct" GDP data. In the absence of real expenditure approach data, Albert Keidel applied various price indices to the different expenditure components, none of which is necessarily fully appropriate for the particular expenditure component.

Wang and Meng's research (2001) is also problematic and lacks reliability. Their calculation of the industrial growth between 1997 and 1999 is not convincing. Wang and Meng's revised evaluation is based on between 141 and 168 industrial commodities. These commodities only occupy about 70 percent of gross industrial output value (GIOV). Why did they choose this number of industrial commodities? Indeed, the other industrial commodities which account for the other 30 percent of GIOV might have helped to achieve fast industrial growth in China during this period from 1997 to 1999. We cannot simply rule out this possibility. If this is the case, how can Wang and Meng argue their commodities sample is quite representative; as they called it "sufficiently responsible" (2001, p. 339)? Wang and Meng's calculation might well understate the real growth of industry and the GDP in China. Moreover, Wang and Meng (2001) further point out that Deng's southern tour in 1992 might have stimulated statistical falsification and manipulation by the local government officers. However, their argument which was based on speculation is unwarranted and without proper evidence to back it up. Deng's speech and southern tour might have led to a few individual local officials to falsify GDP statistics in order to attract foreign investment and stand out in regional economic competition; but I don't believe Deng's tour would have stimulated mass statistical manipulation and embellishment by local officials. As far as Guangdong is concerned, during the early 1990s, regional economic development was indeed impressive; this was not only in the PRDA but also the peripheral region. That can be demonstrated from the counties' statistical figures of industrial, agricultural growth and local tax revenue collection within Guangdong between 1992 and 1994. Through careful cross-year comparison, I did not find that the 1992 economic growth was unusually higher than the GDP growth rates either in the previous year of 1991 or in the following two years of 1993 and 1994. It is important to note that, if the mass manipulation of statistical data done by the local officers in Guangdong is true, it would need the close cooperation of every level of government officers from village, township, and county to municipality in order to avoid disclosure by the media, scholars and ordinary people. We have no hard

evidence of such massive falsification within Guangdong. During the field research, in general, I found that the different levels of statistical bureaus and other government departments in Guangdong generally produce consistent economic and social statistical figures. If mass statistical manipulation and falsification by government officials is the case, this consistency in handling of statistics by the local governments would not be possible. As Holz (2003, p. 155) points out,

None of these articles offers an estimate of the share of enterprises in a particular locality (let alone nation-wide) that resort to data falsification, or on the relative number of local governments that pressure statistical departments into falsifying data.

Chow (1985a, cited in Chow 2005) has also shared this opinion. In fact, the Chinese government has recognized the problems of its statistical data and the many criticisms regarding this issue (The China Daily, 1999, cited in Sinton, 2001). Similarly, Rawski's study (2000) demonstrates that former Chinese Premier, Zhu Rongji openly admitted that there was serious manipulation and embellishment of statistical data. The past statistical manipulation and embellishment during the periods of the Great Leap Forward and Cultural Revolution have brought tough lessons to China. For example, the political campaign of Great Leap Forward during the period between 1958 and 1960 caused massive famine in the southern region (Aijmer and Ho, 2000). As the study conducted by the UNDP (2000) suggests, the Cultural Revolution during the revolutionary period between 1966 and 1976, not only damaged the national economic and social development of China; but also more significantly, the livelihoods of the ordinary Chinese people were seriously harmed by this revolution.

To be fair, both the central and local statistical departments have been making great efforts to improve statistical systems in order to catch up with international standards and enhance the quality and accuracy of Chinese statistical figures. Gustafsson and Shi (2006, p. 138) express their personal perceptions about the staffs of the NBS,

The option has been expressed that the quality of national accounts data in the PRC is probably higher than in many other countries in transition from a planned to a market economy.

In fact, the traditional and backward Soviet-oriented statistical system of data collection, checking and reporting has been significantly changed over the past two decades. Holz (2003) has acknowledged the efforts made by the central government and NBS in improving the statistical system of China. In addition, the central government and Chinese Communist Party have both taken tough actions and strengthened their punishment of falsification of statistical data by local and rural government officers. For example, Holz (2003) cites the case of strong policies by the Chinese Communist Party and individual provinces such as Shaanxi against local data manipulation.

Holz (2003) argues that the statistical inspections at provincial level are more impressive and regular. Some provinces have published relatively detailed statistical regulations based on the Statistical Law of China⁴⁶ and established permanent offices which take

⁴⁶ "The law stipulated that heads of local authorities, departments, or units who altered statistical data,

charge of the task of statistical inspection. For example, in Fujian, there were three statistical inspections implemented during the mid-1990s by the Office of Statistical Inspection of Fujian. In addition, the statistical departments from both central and provincial government have sent their own staff to local statistical offices in order to guarantee the statistical regulations have been fully implemented, and to discover any violations of these regulations in local areas. (Holz, 2003)

Moreover, in order to reduce the over-dependence on statistical figures reported by local statistical personnel, the NBS has started to establish its own rural, urban and enterprises survey teams and use modern technology like the Internet to directly collect comprehensive local statistical data and samples⁴⁷. About one-third of Chinese counties and municipalities have been covered in the working boundaries of the national survey teams. Undoubtedly, this progress is impressive. In addition, about 5,000 key industrial firms have used the Internet channel to submit their economic figures to the NBS since 2001, which represents a significant amount of total sales revenue, tax and profits of Chinese industrial enterprises (Holz, 2003). Currently, there are more and more survey teams directly controlled by the NBS and the establishment of modern statistical systems is under way due to the utilization of various new advanced statistical techniques. The expanding and direct involvement in local statistics data collection by its own survey teams and other actions taken by the NBS demonstrate that the NBS wants to reduce its high dependency on local bureaus and eliminate statistical collection and data reporting by the local governments of China. The Chinese authority has a sufficiently good reason to develop a modern, transparent, and high-standard statistical system and practice; it is in China's national interest to do so. As Chow (2005, p. 1-2) argues,

There is no incentive for the Premier [of China] to lie not only because his report is under the scrutiny of the entire world but because using fabricated data can only lead to confusion of all government officials using such data for their work in economic planning and development. (Bold content added)

Relative to the Maoist era of economic planning, establishment of a modern market economy in China since 1978 has naturally forced the NBS and its lower local statistical departments to take actions in transforming the traditional and backward statistical reporting system in China. Without high quality and reliable statistical system, the Chinese government would find it extremely difficult to set up proper plans for long-term economic and social development; it would be also impossible to effectively supervise various economic activities under modern market economy. As Wang and Hu (1999, p. 224) argue, "In the post-Mao era, accurate economic data became vital for macroeconomic management." On the one hand, the Chinese media, research scholars and other Chinese people have obtained more and more information through various channels such as Internet, radio, television and newspaper. On the other hand, with rapid economic development and greatly improved communication with the outside world, the Chinese people are increasingly demanding high quality statistical data. These factors

fabricated statistical data, or compelled or prompted statistics institutions or statisticians to tamper with or fabricate statistical data would be given administrative sanctions. If the case constituted a crime, persons involved could be investigated for criminal responsibility." (Statistical Law of People's Republic of China, 1983a, chap. V. article. 26, cited in Xue, 2004, p. 93)

⁴⁷ As far as the statistical system in China is concerned, Zheng (2001) offers a detailed discussion regarding the three vertical channels of data collection.

have also forced the Chinese authorities to take action and transform the traditional Soviet-oriented statistical reporting practice and provide reliable statistical data. Even Rawski (2000) admits that China is on the way to becoming an open society and important aspects of economic and social history: like the pre-reform period, cannot be kept secret.

Some critics (e.g. Maddison, 2006) argue that it is impossible to collect and produce generally reliable and accurate statistical data through a reporting system which is mainly hierarchical. Maddison's criticism is partly right. The Chinese economic system itself had been almost wholly destroyed. The small amount of statistical data published by the central statistical department largely depended on local data collection. However, this situation has been significantly improved during the reform period. The NBS has largely eliminated collection and calculation of national statistical data based on the simple sum calculation of raw statistical figures collected by the local regions. For example, on several occasions, the NBS has openly rejected the subtotal GDP figures collected by the provincial-level statistical authority, such as the 1997/1998 GDP growth rate. (Rawski, 2000)

Moreover, in August 2006, the NBS again rejected the provincial GDP's growth rate (12 percent) for the first half year of 2006 due to suspicion of local statistical embellishment; by contrast, the NBS maintains that 10.9 percent would be a more reliable growth figure for the first half year of 2006. The difference of total GDP announced by the central and provincial statistical departments was significant; the number was 804 billion *yuan* (Southern Metropolitan Newspaper⁴⁸, 15th August 2006). Holz (2002 and 2003, cited in Holz, 2005) has shared this opinion that the NBS has not automatically accepted and utilized the raw statistical data collected by local statistics bureaus. Similarly, Wang and Hu (1999) also stress this point.

The statistical system and practice in China are far from perfect or desirable, and despite tough actions by the central government and the NBS to eliminate statistical manipulation, there is still room for improvement. Nevertheless, it is unfair to turn a blind eye to the impressive progress that has been made or to completely reject Chinese GDP statistics. It is also important to recognize the difficulties and massive task involved in transforming China's backward statistical system of the planning period⁴⁹ into an advanced statistical system during the market economy era.

In terms of developing the Chinese statistical system and departments, good progress has been made during the reform period. Only thirty years ago, China was still suffering from the huge disasters of Cultural Revolution. These kinds of political campaigns almost

⁴⁸ Southern Metropolitan Newspaper (*Nanfang Dushi Bao*) which is part of the South Daily Newspaper Group Ltd (SDNG) (*Nanfang Baoye Jituan*), The SDNG is under the direct supervision of the Propaganda Department of Guangdong. Southern Metropolitan Newspaper is a very popular newspaper for the government officers, academic scholars and ordinary Cantonese people; its credibility has been widely recognized by the academic community and media within Guangdong. Average daily sales of this newspaper are about 700,000.

⁴⁹ Keidel's study (2001) has provided a detailed discussion with regard to the Soviet-oriented statistical system and practice of China from Mao's era to the early 1990s. Keidel has also pointed out the shortcomings and limitation of NMP and MPS, such as exclusion of the majority of output of tertiary industry. The study by the World Bank (1992, cited in Keidel, 2001) suggests that the NMP approach has suffered from problems of improper estimation of economic activities of net exports, and miscalculation of some transactional activities.

completely destroyed the Chinese political, social and economic system. Unfortunately, the statistical system is not an exception. The Chinese statistical system was almost completely abolished by the rebellion faction (*zaofanpai*) and the Red Guard (*hongweibing*). As suggested earlier by Maddison, the statistical departments could not be properly run; many skilled statistical personnel were dismissed; and statistics students in the universities were forced to give up their studies and go to the countryside. However, during the reform period, China has been making significant progress regarding modernization of the statistical system. As Maddison (1998, p. 57, bold content added) honestly points out, “Since 1978, the situation [of Chinese statistical system] has improved greatly, the accounts are more transparent, coverage and classification more or less conform to Western concepts.” He further stresses that China has been making impressive economic achievements during the reform period, referred to him as “China enjoyed supergrowth” by him (1998, p. 55). Rawski (2000) and Xue (2004) also acknowledge the progress made by the statistics departments and the efforts put in by statistical staff in China during the reform period. Wang and Meng’s (2001) bias-oriented criticism is far more sceptical towards the significant work done by the NBS.

In contrast to its past history and inherited traditional statistical system, it is important to note that the comprehensive modernization of statistical systems and practice is not an easy task for China. As Rawski (2000, p. 17) points out, “China’s immense size multiplies the task of creating or revising systems of data collection. It also multiplies the scale of reform achievements in improving the measurement of economic activity.” In addition, Holz (2004) has also recognized the difficulties of collecting and analyzing accurate statistics in China during its period of economic development. In comparison to the Maoist era, China is moving toward an open and transparent market economic system. It means that big social and economic issues cannot remain hidden from the public as in the past. Moreover, without question, China has been transforming into a more open and democratic society. Rawski and Xiao’s study (2001) has shared this opinion. For example, today, many sensitive economic topics which were only reluctantly talked about or even recognized in the pre-reform period have been raised and discussed by scholars, media and governments; such as the unemployment rate, regional inequality, poverty, pensions and stock market. Rawski’s study (2000) suggests that Chinese economists now frequently discuss current urban unemployment in China. Interestingly, some scholars suggest that statistical data published by the NBS might have underestimated the actual growth of the Chinese economy. Fan Gang (cited in Holz, 2003) argues two sources for this under-evaluation of Chinese statistics: the rich and poor regions in China.

8.2.4 Official Statistical Data of Guangdong: Reliability and Accuracy

In this research project, I refer to statistical data within Guangdong, ranging from township, county and municipality, to province. In addition, I have cited various aspects of local statistical figures, such as population, land area, industrial growth, agriculture, weather. How accurate are Guangdong’s statistics in contrast to central and other provincial statistics, in particular, the county-level statistical data?

During my fieldwork research, I did experience many difficulties in collecting and analyzing local statistical data; this is especially true of the least developed areas like Wuhua. The collection and analysis of local statistical data done by official departments in these economically poor regions are notably poor. The data in many important areas, such as life expectancy at birth and educational development, have not yet been kept or

reported. This is perhaps due to lack of financial resources and personnel in these official statistics departments. Moreover, data inconsistency makes the cross-county and time-series comparison of regional income disparities within Guangdong difficult to carry out. In addition to this, the discrepancy of statistical data between the central and provincial statistics bureaus is sometimes quite remarkable, which raises doubts as to their credibility and accuracy. I took quite a long time to judge, analyze and resolve the problem regarding data inconsistency in order to provide a reliable overview of spatial inequality. This problem of data inconsistency is concentrated on the poor counties within the peripheral region; the discrepancies of GDP statistics between Guangdong Yearbook and Guangdong Statistical Yearbook are notable in some years from 1998, for example, in Puning and Shantou. This issue sometimes makes me seriously concerned about the accuracy of statistics in Guangdong. Director of the Guangdong Statistics Bureau (GSB) – Bo Xinmin, admits that,

As far as the statistical data of Guangdong is concerned, the performance-motivated statistics figures in some regions within Guangdong still exist. *This problem has not only affected the objectiveness of statistical figures, but also this issue has badly influenced the accuracy of statistical data, and caused the data inconsistency between the central and provincial statistical departments.*

(Southern Metropolitan Newspaper, 21st March, 2007, italic emphasis added)

More significantly, in Guangdong, the increasingly serious problems of smuggling and corruption have brought additional concerns over statistics quality. Smuggling not only brings heavy loss of total national and local tax revenue, but also could reduce accuracy of local GDP statistical data. This is particularly relevant to the total amount of exports and imports. Indeed, the general accuracy of GDP figures could largely depend on the extent of smuggling. Cheng's study (2000) has suggested the serious impact of smuggling on various coastal regions within Guangdong; this is especially true of Zhanjiang and Shantou.

The rampant corruption by local governmental officers is another serious issue affecting the quality of local GDP figures in Guangdong. During the 'reform and open-door' period, corruption and rent-seeking behaviour have become severe. If corruption is widespread within a region, how can the people believe this region can produce reliable GDP statistics? In the early study conducted by Vogel (1989), he highlights the prominent issue of corruption in Guangdong. Furthermore, Cheng (1998; 2000) argues that Guangdong is particularly the most corrupt province in China. Under these circumstances, corruption in Guangdong needs to be taken into account when analyzing local statistical quality.

It is certainly the case that local statistical data in Guangdong has been facing a quality problem. This is particularly serious in the underdeveloped and rural regions. Incidences of serious manipulation of statistics have been steadily increasing in recent years. For example, in 2004, 232 individual cases were investigated and confirmed by the GSB; cases of serious manipulation were up to 267 in 2005, while equivalent cases further jumped to 288 in the year 2006. Although tough actions and measures have been taken to deal with local statistical manipulation, the figures suggest that cases had increased more than 24 percent during the three years from 2004 to 2006. In addition, there are only 87 counties and districts within Guangdong which can be fully covered by the routine

inspection teams of the GBS; by 2004, around 20 counties and districts had still not been covered by the statistical law inspection and enforcement conducted by the GBS (The data was collected from the GSB, available on the official website of the GSB, accessed in April 2007). During our field interview conversation, in reply to question of statistical quality in Guangdong, the Vice-Governor of Guangdong stressed that the provincial government has paid special attention to this issue. Fast economic development and export-oriented trade in Guangdong demand accurate data, if the government is to supervise economic and industrial development effectively. He further states that,

We do not disturb the daily working affairs conducted by the Guangdong Statistics Bureau (GSB), and we do not put our influence to the statistical staffs regarding data collection and reporting; we respect the expertise and integrity of the professional statistics personnel in the GSB. (Field Interview 1, September 2006)

Huang's argument might be too optimistic and overestimate the reliability of Guangdong's statistical data, but it reflects that the GSB has implemented its duties by checking independently with less administrative influence from the provincial government. Further incentive to improve efficiency and reliability may come from the fact that in order to attract more foreign capital and advanced technology, the provincial government of Guangdong has to change its traditional and backward economic system and follow international standards. Establishing a modern, reliable and accurate statistical reporting system is one of the important tasks. During a field interview with an anonymous member of senior statistical staff from the GSB, she stated that,

We are doing our best to improve the statistics quality of Guangdong. The increasing pressure from the up-level government has forced us to take actions to enhance the quality and accuracy of our statistical data. One thing I want to highlight in here is this: even if our statistical data can not be precisely accurate like the specific annual accounting report of the individual enterprise, our statistical data are generally reliable and valid. (Field Interview 9, October 2006)

Holz seems to also share this opinion. Holz (2003) points out that misinterpretation of particular data is the main criticism of Chinese statistics. He suggests that severe data falsification in the provincial and central statistics departments is almost impossible.

In addition, the GSB and other government organizations in Guangdong have taken further action to deal with the issue of statistical manipulation and embellishment, and to ensure the quality of local statistical data. The GSB has been reinforcing inspection and investigation in the lower statistical departments according to the Regulation of Statistics Law Inspection and Enforcement of China⁵⁰. Moreover, the GSB has taken the following

⁵⁰ The Regulation of Statistics Law Inspection and Enforcement of China came into force on 11th May 2001. It has set up precise rules on the handling procedures of statistical data inspection and investigation by the local statistics bureaus of China; this regulation has also established the specific investigation procedures on the violation of statistics law and regulation for the statistics bureaus of China. In addition, it also designates the detailed punishments on breaking Statistics Law and Regulation of China by the government and commercial organizations, and any violent behaviour toward the statistics personnel (the GSB, available on the official website of Guangdong Statistical

measures to check and ensure the accuracy of statistical data collected by local statistical personnel.

- First, extending and strengthening investigation and checking of the statistical figures reported by local enterprises: focusing in particular on the growth rate and total value of industrial output. In particular, the GSB pays more attention to the enterprises which reported obviously high growth rate of industrial output which was inconsistent with statistical indicators.
- Secondly, strengthening the investigation and supervision of local statistical bureaus. The working group of GSB (*gongzuo zu*) has made frequent inspections of local statistical bureaus; trying to ensure that the proper methods of statistical collection, checking, reporting, storing and management regulated by the Statistical Law of China, have been fully implemented by the local statistical bureaus within Guangdong.
- Thirdly, strengthening data collection skills and enhancing the ability of local personnel through training courses. In particular, special attention has been given to newly recruited, local statistical personnel. In order to ensure the quality of industrial statistical data, the GSB hopes to ensure that all of these new local staff fully understands the statistical reporting system, the definition of statistical indicators and the relevant calculation methods.

(Source: GSB, 2006)

In addition, the municipal statistical departments, in Guangdong have also increasingly strengthened their duties of statistical investigation and inspection, and taken tough measures on dealing with illegal activities of statistical manipulation and falsification. For example, the Statistical Bureau of Guangzhou publicly reported illegal activities of statistical manipulation and embellishment in 2004. There were 45 individual cases of statistical manipulation that had been investigated and confirmed by the Statistics Bureau of Guangzhou and 12 cases were released to local media. In 2006, the number of a investigated and confirmed cases had risen to 94 and 11 organizations and enterprises were listed in the local newspaper; in particular, the name of a local government was included; the Township Government of Taiping, Conghua City. All of these organizations were severely punished and the people directly involved were fired. In addition, the Municipal Government of Qingyuan publicly criticized the improper conduct by its three township-level governments; and also requested them to take actions to improve the system of statistical data collection and reporting (GSB, 2007).

During the field research in Guangdong, I found that the various local statistical data collected and utilized in the different government organizations within Guangdong were generally consistent and agreed with the published statistical figures without any significant discrepancy. If the local authorities had deliberately manipulated and embellished the statistical data, it would have been difficult to keep and publish consistent data all the time. Holz (2003, p. 153) points out that,

Widespread data falsification would surely lead some employees of the statistical departments at some point to reveal the existence of a second set of data. Various internal publications that are occasionally available to researchers largely match the published data.

Bureau). The detailed rules of this regulation can be obtained from the official website of the NBS.

In my view, while there is some manipulation of data in Guangdong, the problem is on a small scale and is mainly concentrated in the least developed regions. More significantly, the general trend of dramatic economic development in Guangdong during the reform period has not been fundamentally affected. As far as this study is concerned, the time-series regional economic inequality within Guangdong is the main focus of investigation. Problematic local data regarding the GDP growth and GDP per head will not seriously challenge the development tendency of spatial disparity during the reform period. In that sense, to some extent, the data problem in Guangdong is not severe, and it can be tolerated and accepted by this study.

Guangdong is a reform pioneer and one of the earliest provinces in China to engage in foreign trade and exports and to open up to foreign investors. Moreover, Guangdong is highly complementary with Hong Kong and Macao in terms of people and information communications, commerce, trade, and finance; these advantages established the ideal foundation for Guangdong to access and obtain the latest global information, and adapt to common international practice and advanced economic systems. Vogel's study (1989) highlights the close economic correlation, and the transfer of advanced management skills, technology and information communication between Guangdong and Hong Kong.

With the rapid economic growth and increasing local tax revenue, relative to other provinces of China, Guangdong could put more public investment into developing comprehensive and high quality statistical systems during the 'reform and open-door' period. Geographical proximity to Hong Kong means that Guangdong has the competitive advantage of faster acquisition and understanding of advanced international statistical practice. Personally, I believe that Guangdong's statistical data are generally reliable and accurate: at least no worse than in the other regions of China. If the national statistics are generally reliable, the utilization of Guangdong's statistical data should be also safe in academic research. Of course, careful judgment and selection of local statistical data is needed and awareness of the potential limitations of local statistical figures in Guangdong should be also maintained throughout the research project.

8.3 Conclusion

It is reasonable to argue that both the GDP statistical data in China and Guangdong have certain problems with regard to quality and reliability; data discrepancy between the central and local statistics departments is a particular concern. Without question, in order to fully improve its systems and practices of statistical collection and analysis, and ensure the quality and accuracy of its data, China still has a long way to go. China not only needs to clearly and consistently explain definitions and concepts of statistical sectors like the secondary and tertiary industries, but also to deal with the data inconsistency problem. Although both the central and local statistical bureaus have put more effort into tackling violations of the Statistics Law of China and other regulations, data exaggeration and manipulation by local officers still exists. As discussed earlier, the dramatically increasing number of violations of statistics law and regulation in Guangzhou and the other regions in Guangdong have demonstrated this point. Moreover, the capability of statistical data collection by the NBS is still questioned. The revision statistics figures of 2004 Economic Census reveal this fact.

Nevertheless, both the NBS and GBS have made sincere efforts in the modernization of traditional statistical systems and practice in both Guangdong and China as a whole; the improvement in the Chinese statistical system is impressive. Significantly, the Chinese GDP and other statistical figures have correctly reflected the fast economic and social development of China during the 'reform and open-door' period. It is generally safe to use Chinese GDP statistical data in academic research. As far as this study is concerned, rather than the single year, time-series measurement and analysis of regional economic inequality is the key. The problems of Chinese statistics quality are not sufficient to change the measurement outcome of regional inequality in Guangdong.

However, scholars and policymakers need to approach Chinese statistical data with care and make necessary declarations in their research. As Chow (2005, p. 18) points out, "Chinese official data are by and large reliable enough for serious scholars to perform econometric analysis to produce useful results although some data may contain errors. Any serious scholar using these data, as in using any other data needs to exercise caution in his research." The accuracy of Chinese statistical data is still debatable; however, the general credibility of the Chinese statistical system, the integrity of the NBS, and also the important work done by the NBS and its lower provincial statistics departments should be properly recognized.

In conclusion, there are numerous criticisms regarding the quality and reliability of Chinese statistics; but, many of these criticisms suffer from their own limitations. They are not able to seriously challenge the fact that Chinese economic development has dramatically increased over the past two decades. These criticisms are also unable to change the general estimation of economic and industrial growth in China conducted by the NBS. More significantly, with the 9 percent economic growth claimed by the officials, or even at only 5.7 percent or so argued by the critics, in contrast to other countries, the economic growth rate of China has certainly been outstanding during the reform period.