Regional Welfare Disparities and Regional Economic Growth in Vietnam

Promotor: Prof. Dr. W.J.M. Heijman
Hoogleraar Regionale Economie, Leerstoelgroep Economie van Consumenten en Huishoudens

Co-promotor: Dr. J.A.C. van Ophem
Universitair Hoofddocent, Leerstoelgroep Economie van Consumenten en Huishoudens

Promotiecommissie:
Prof. Dr. Ir J.D. van der Ploeg (Wageningen Universiteit)
Prof. Dr. H. Visser (Vrije Universiteit, Amsterdam)
Prof. Dr. A. Kuyvenhoven (Wageningen Universiteit)
Prof. Dr. J. van Dijk (Rijksuniversiteit Groningen)

Dit onderzoek is uitgevoerd binnen de "Mansholt Graduate School of Social Sciences"

Regional Welfare Disparities and Regional Economic Growth in Vietnam

Nguyen Huy Hoang

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit
prof. dr. M.J. Kropff,
in het openbaar te verdedigen
op dinsdag 17 maart 2009
des namiddags te vier uur in de Aula

Nguyen Huy Hoang (2009)

Regional Welfare Disparities and Regional Economic Growth in Vietnam

Hoang, N.H.
PhD thesis Wageningen University (2009). ‒With ref.
‒With summaries in English and Dutch.
ISBN: 978-90-8585-319-0

Preface

Of the past five years, since my arrival to Wageningen to start my PhD research, especially the first three and half years were not always a smooth sailing. I had to catch up to courses, construct empirical models and improve my English writing skills. Sometimes, finishing the thesis seemed far away. Apart from these problems, the PhD period has been an exciting one, with many opportunities to travel, talk to interesting people, make new friends, develop new skills and, above all, I have learnt a lot. Many times it was really nice and the job would not have been so pleasing without the interaction with people to whom I am very grateful.

First of all, I need to show my great appreciation to my promoters Wim Heijman and Johan van Ophem, who have both been vast sources of inspiration and positive feedback. Wim and Johan, I thank you both very much for your hard work reading my chapters and giving me valuable suggestions and corrections, as well as the encouragement and enthusiasm you shared with me over the years, especially Johan, you only took the job one year ago, but the help you devoted to my work was huge and invaluable. I would also like to show my sincere gratitude to Henk Folmer for his cooperation as my promoter in the first four years of my research before Johan came to scene. Henk, thank you very much for the multidimensional approach to welfare inequality you introduced to me and the supervision you gave me initially.

My special thanks go to all members of the Economics of Consumers and Households Group of Wageningen University. Most importantly, Annelies Coppelmans and Margaret van Wissen, who always helped me when I needed that. Other staff of the group: Gerrit Antonides and Pierre van Mouche. My former and current PhD colleagues and friends of the group: Judith Cornelisse, Masayo Igata, Muyeye Chambwera, Pius Odunga, Subrata Dutta, Hirut Bekele, Leonie Cramer and Romanus Dimoso. Thank you all for your company during the time I was there in the group. I have enjoyed the frequent discussions we had about every aspect during coffee breaks or whenever we gathered together. I am also grateful to Ms. Dinie Verbeek for her help making the thesis ready for printing.

Thanks are also extended to my fellow doctoral colleagues from other departments of Wageningen University, my fellow doctoral Vietnamese in Wageningen University and many others who were with me during the time I was there. I have enjoyed the discussions we had and travels with you during our stay in Wageningen.

I would also like to thank my colleagues in Vietnam: Nguyen Huy Thanh for your various formal administrative procedures, and Nguyen Van Ha, Nguyen Ngoc Lan, Vu Thi Van Anh, Le Phuong Hoa, Quang Thi Ngoc Huyen and Nguyen Thuong Huyen. You were always sources of fun to help easing my frustration and tension I faced during the time I was constructing and estimating empirical models for my thesis, and writing chapters. Thanks also go to Vu Tuan Anh from the National Institute of Economics for his cooperation and assistance.

I would also like to gratefully acknowledge the financial support I have received from the Ministry of Education and Training (MOET), Government of Vietnam. Many thanks also go to individuals (Pham Manh Hung and Le Anh Kiem) and institutions (Department of Socio-environment Statistics and Department of National Accounts of the General Statistics Office of Vietnam - GSO) that helped me during the data collection phase in Vietnam. Without your help, this thesis would not have been possible.

Last but not least, I am deeply indebted to my family and most beloved people. My daughter, Bao Phuong, and my son, Duc Nguyen, have been a source of inspiration for me during the PhD period. It seems to me that my late father has always followed my every step. My mother was always behind me. My wife has sacrificed a lot to help me looking after our children and taking care of our family when I was away, and she has been a huge source of support for me during the past five years. I am really grateful to her for all kinds of support and sacrifices. My brothers, sisters, nephews, nieces and my most beloved friend, Nguyen Trong Minh, who have always loved me and taken care of me and my children over the past years. This thesis is partly devoted to them.

Once again, I thank you all!

Nguyen Huy Hoang

Table of Contents

List of Tables

List of Figures

Chapter one
Introduction

The study of regional welfare inequality and its linkage to regional economic growth have drawn a lot of attention in the literature of development economics. However, the study of regional welfare inequality and its relation with economic growth in Vietnam has not received proper attention so far. In addition, most studies on inequality focus on measuring and analyzing income or expenditure inequality by using data from various rounds of household living standard surveys, and by applying either the Gini coefficient or the Theil's index, using the traditional method of inequality analysis, the single-indicator approach. Furthermore, studies on the relationship between regional income inequality and regional growth, growth patterns and convergence in income among provinces of Vietnam, as well as spatial dependence and role of space in Vietnam, are almost absent. Since the implementation of the economic reform called *doi moi* in 1986, which transformed the country from a centrally planned system into a market-oriented economy, the distribution of welfare and the patterns of economic growth have considerably changed in the country. Studies by the World Bank and the General Statistical Office (GSO) of Vietnam (various years) and several researches by, among others, Minot *et al.* (2003); Heltberg (2003), Liu (2001) and Dollar *et al.* (1998) have demonstrated that the economic reform in Vietnam has led to substantial movements in relative prices and wages that affect welfare inequality in general and regional welfare inequality in particular, and has led to an increase in income disparities between social groups and between regions. Also, these studies reveal that the regional single-indicator inequality in health and education over time has changed. However, not any study so far has taken these components of welfare as a multiple indicator composition of welfare.

1.1. The objectives of the study

Regional economic growth and regional welfare inequality are two important and sensitive topics in current development studies. The study of welfare inequality is very important both in literature and empirically. In the literature on welfare inequality, two approaches are often used to explore the status of welfare distribution, namely the single-dimensional approach and multidimensional approach. The former has been widely used from the beginning of studying welfare inequality and is sometimes referred to as the traditional method of measuring and analyzing welfare inequality, while the latter is used when researchers recognize that a welfare component consists of not only one dimension or indicator (for example, income or expenditure), but also of more than one dimension or indicator (such as health, education or other basic needs). Thus, without the multiple indicators in one welfare component, there would be no need for the multidimensional approach. When such multiple indicators exist in a welfare component, we need to find a composite measure that is a function of the welfare indicators of concern to us. Therefore, the multidimensional approach is applied to study welfare inequality when multiple indicators are included in the welfare inequality measuring process. Recently, the study of welfare inequality using the single approach has been criticized for inadequately and insufficiently reflecting the status of welfare inequality, and usually causing misinterpretation (underestimation or overestimation) of the welfare inequality situation (Arrow, 1971; Sen 1970, 1973, 1977; Kolm, 1977; Atkinson and Bourguignon, 1982; Sen and Foster, 1982; and Maasoumi, 1979, 1986). In addition, welfare itself is multidimensional, because the welfare components comprise multiple aspects or indicators of welfare, such as the human capital elements as education and health status. The education component consists of several indicators that represent the education status, such as enrolment rates, education facilities; or the health component also consists of many indicators, such as number of doctors per thousand population, number of pharmacists per thousand population and number of hospital beds per thousand population. Thus, using the single dimensional approach to consider inequality in the distribution of welfare could lead to a wrong interpretation of the welfare distribution of such welfare.

For the reasons mentioned above we will apply both the single indicator approach and multiple indicator approach to study the welfare distribution when it is appropriate. For example, for measuring and analysis of regional inequality of health (chapter four), and measuring and analysis of regional inequality of education (chapter five), the multidimensional approach and a function of the welfare indicators or composite measure are used, while the measurement and analysis of regional income inequality, the single-dimensional method is used (chapters six and seven).

Regional welfare inequality is affected by economic growth. As pointed out by Kuznets (1955) and his successors in this field, there is a positive relationship between inequality and growth in the initial period of development, but the relationship turns to negative at a later stage of development. Moreover, space is an important factor that affects regional growth, regional welfare inequality and the relationship between regional growth and regional income inequality (Anselin, 1988; Chen and Fleisher, 1996).

Despite the fact that income inequality in particular and welfare inequality in general have increased in Vietnam, there are several issues related to inequality in the welfare distribution that have remained unresolved until now. The measurement and analysis of the multidimensional regional welfare inequality have not been conducted in Vietnam so far. Thus, in the first place this study will aim at measuring, analyzing and exploring the magnitude (size) and trend of multidimensional regional welfare inequality in Vietnam. The welfare components to be taken into the analysis are health and education. In addition, the measurement and analysis of multidimensional regional inequality in education and health, based on the composite index of each welfare component, regional rankings for health and education are constructed, and, for the aim of policy implication, regions are classified as most favored, medium favored and least favored with respect to the composite index of the welfare components as health or education. Parallel with measuring and analyzing regional welfare inequality for health and education, and ranking education and health with respect to their respective composite index, the study will also focus on the role of space or spatial dependence. For the health component of welfare, we attempt to measure and analyze regional inequality under two categories: contiguity and non-contiguity, as we assume that health facilities in one province are available not only to its own residents, but also to those living in adjacent provinces, which may affect the health status of the people in adjacent provinces. These separate analyses are not applied for the case of education component of welfare, because there is a mandatory mechanism applied to education, which means that pupils are not freely allowed to get education in other locations than in the place they are registered to live, except university education and technical training.

In the second place, the subject of regional income inequality within a particular country, the relationship between income inequality and economic growth in general, and between regional income inequality and regional economic growth in particular, has been an issue that has received much attention and dispute among economic researchers (in the past as well as at present). This issue has been brought to the agenda of development economics since the study by Kuznets (1955), which found a positive relationship between income inequality and economic growth in the initial period of economic growth, and a negative relationship at a later stage of development. Recently, many studies on the relationship between regional economic growth and regional income inequality have been conducted for the developing and transformed or transitional economies, such as China and Russia, which have found both a positive and negative relationship between regional economic growth and regional income inequality, see among others, Solanko (2003), Ahrend (2002), Chen and Fleisher (1996), Weeks and Yao (2003), and Bradshaw and Vartapetov (2003). However, studies on the relationship between regional income inequality and regional economic growth in Vietnam have almost been absent so far. To investigate such a relationship, our study will start with the measurement of regional income inequality in Vietnam. Then we go a step further to consider and analyze the relationship between regional income inequality and regional economic growth, by using regression models and tests for the regional effects on economic growth and income inequality in Vietnam. To do so, we apply the one-stage Theil's decomposition method to measure regional income inequality. Then we use the linear regression model to explore the relationship between regional inequality and regional economic growth with the inclusion of several control and dummy variables that affect both the inequality and growth rate of income.

In the third and final place, the present study will go in detail into the issue commonly raised in development economics literature. That is the growth pattern of regional income in Vietnam. To investigate the pattern of regional income growth, the study will employ various tests for convergence in per capita GDP among provinces in Vietnam. To do so, the growth model is used and several tests for convergence, such as Moran I test, Lagrange Multiplier test and Robust Lagrange Multiplier test, as well as the Unit root test are employed to detect whether there is a sign of convergence in the growth pattern of regional per capita income in Vietnam.

In short, the main purposes of the analysis in this study are multiple. Firstly, the study aims to gain a better understanding of the magnitude and trend of regional welfare inequality in health and education, and regional income inequality. The study also classifies regions with most favored and least favored with respect to the composite index of health and education for proper policy intervention, in order to improve the health and education system in Vietnam. Secondly, the study helps to get a better understanding of the contiguity effects in the case of health facilities and their impact on health status among regions in Vietnam. As argued above, there is mandatory education applied to students, so the contiguity effects will not be taken into consideration for the education component. The reason behind the selection of only health (in chapter four), education (in chapter five) and income (in chapter six) for measurement and analysis of regional inequality lies in the availability of the data. We only find systematic data on health, education and per capita GDP at the provincial level that are sufficient and trustworthy for analysis of inequality in health, education and per capita GDP.

Parallel with the measurement and analysis of multidimensional regional inequality in health and education, the study also examines regional inequality in per capita GDP and analyzes the relationship between regional inequality in per capita GDP and regional growth rate in order to help understand the causal relationship between growth and inequality. Lastly, the study goes a step further to analyze the pattern of regional economic growth by examining whether the regional growth is converging together or it is diverging from each other. This section would provide knowledge about the trends and patterns of the regional economic growth in Vietnam.

1.2. The research questions of the study

Regional welfare inequality in Vietnam is changing. Many studies on the distribution of welfare have been conducted in the country so far, but these studies focus only on the distribution of income or expenditure, or distribution of each indicator of a welfare component, such as health or education. Studies that take the multiple aspects of a welfare component into consideration, using composite function of welfare indicators (variables) are almost absent. In addition, one important feature, the long-term topic of welfare inequality, its magnitude and trends, and their relationship with economic growth, has not been fully explored. Another important topic that is examined in this study is the pattern of growth over the period of high economic growth from 1990 to 2006, which also needs to be taken into account. Thus, there are still many questions relevant to multidimensional and uni-dimensional patterns of welfare inequality, as well as the role of space or spatial dependence in the relationship between growth and income inequality and pattern of growth that remain unanswered. Thus, the present study aims to seek answers to the following questions:

- What is the size of regional multidimensional inequality in health and education?
- How has the regional multidimensional inequality in health and education evolved?
- What does the regional profile based on the rankings of composite index with respect to education and health look like?
- What is the role of contiguity effects on the health status of the people or does a spatial dependence among health status and health facilities in the country exist?
- What are the magnitudes and trends of regional inequality in per capita GDP?
- What is the relationship between regional inequality in per capita GDP and regional growth in Vietnam? What is the role of space in this relationship?
- What is the regional pattern of economic growth? Does the regional per capita GDP converge to or diverge from one another?

To find answers to the above-raised questions, this study considers the case of Vietnam. The study is based on the provincial data for the 61 provinces in eight administrative regions in the country (see Appendix 1 of chapter two). The study is carried out for these 61 provinces, and considers each province a region, whereas Vietnam is considered a global entity. The thesis is constructed and organized in eight chapters.

1.3. The approaches of the study

As already mentioned, this study concentrates on measuring and analyzing multidimensional regional inequality of several non-income welfare components, such as health and education, by using Theil's second index of inequality. This makes the first discerning characteristics of the present study in the sense that it deals with the regional inequality that almost no or only few previous studies have looked at. Other components of welfare such as, for example, basic needs and quality of life will not be included in this analysis. The reason for this choice lies in the absence of data on basic needs and quality of life in Vietnam (see chapter three, chapter four and chapter five for a more detailed discussion). In addition, in the analysis of each welfare component as education or health, several indicators are not included in the analysis such as, for example, higher education facilities at university or equivalent level (number of high education institutions and number of high-education teachers) and higher enrolment rates in the education component of welfare or calorie intake in the health component of welfare. The reason for omitting these variables is the lack of systematic data on these indicators in Vietnam. The present study also takes into account the contiguity effects for the case of health, in order to consider the spillover caused by variables that represent health facilities that directly affect health status in neighboring regions as well.

Another important difference with previous studies on the subject of welfare inequality lies in the approach that is used for analyzing welfare inequality. The commonly used method is applying the single dimensional approach to consider welfare inequality. In the present study, we employ the multidimensional method to measure regional inequality in education and health, by using Theil's index, which can be decomposed into between- and within-components of inequality. The salient feature of the approach used in the present study is the use of the aggregate function developed by Maasoumi (1986) to compute a composite index. This feature makes the approach more trustworthy and unique in the measurement and analysis of welfare inequality for Vietnam.

The measure of regional income inequality that will be used in the analysis of regional disparities is also the Theil's index. We use the one-stage decomposition technique to find answers to the question of regional income inequality, trends and magnitude in the provinces of Vietnam. The measure of income inequality in this study is different from the measure of regional welfare inequality for health and education. For the former, there is a single indicator as per capita income is considered, while for the latter, there are two or more variables taken together for the calculation of the composite index e.g. Theil's second measure. After measuring regional income inequality, another approach is used to explore and analyze the relationship between regional inequality and regional growth rate using the regression model method.

Lastly, another framework is employed and used in chapter seven of the present study to examine patterns of economic growth in Vietnam and the role that space plays in the growth process of the country, by using growth models and several tests of regional convergence of the per capita GDP, such as Moran I test, Lagrange Multiplier test, Robust Lagrange Multiplier test and the unit root test. The framework used in this chapter is developed from the approach that has been widely used in the literature of development economics in order to examine growth patterns and to test for possible income convergence among countries or regions in the world or among regions within a country in both developed and developing and transitional countries.

1.4. Outline of the study

After the introduction presented in chapter one, we elaborate on and discuss macroeconomic reforms and changes in welfare in Vietnam in chapter two. After a long economic recession in the postwar period (war with United States of America or the Vietnam War) during the late 1970s and early 1980s, the country has started the economic reform or renovation, called *doi moi* in the Vietnamese language, in 1986. Thus, this chapter at first elaborates on the process of economic development before the *doi moi* for the period from 1975 to 1985. But before doing so, the chapter provides readers an overview of the geography and demography of Vietnam and its regions. Then the chapter goes on to discuss the key characteristics of the reform policy, achievements and its impact on macroeconomic performance. This chapter also provides in detail the macroeconomic changes in various five-year plans, as the government of Vietnam had set its strategy for each five-year plan of economic development. After discussing and elaborating the macroeconomic performance, chapter two goes on to elaborate on the changes in welfare in the country. In detail, the chapter discusses the improvement in welfare and living standard of the people and the reduction of poverty brought about by the *doi moi*. Also, the chapter discusses trends, patterns, dynamics and size of income inequality in Vietnam over time, stretching from 1975 to recent years.

In chapter three, the method of the multidimensional approach to regional welfare inequality is discussed and presented. In this chapter, we provide the basic background for using Theil's second index and a brief review of literature that deals with welfare inequality in general and multidimensional regional welfare inequality in particular. This chapter presents the theoretical model to compute Theil's second index used by Akita (2003) and the aggregate function for composite index calculation developed by Maasoumi (1986) and used by other researchers, including Quadrado (2001). Measurement models are also extensively discussed and presented in this chapter. The chapter elaborates on the method to measure weights attached to each indicator that is composed of a composite index. The method is developed and used by Flury (1988) and called principal component analysis (PCA) or partial common principal component analysis (PCPC).

In the empirical analysis, we start with measuring and analyzing multidimensional regional inequality for the health component of welfare in chapter four, using the method developed in chapter three, for health facilities and health status separately. For health facilities, the chapter reports the results for two cases: contiguity and non-contiguity in order to examine the spillover effects in health facilities among provinces in Vietnam. Besides, the chapter examines the spatial patterns between health facilities and health status by both visualization and a spatial econometric model developed by Anselin (1988) and LeSage (1999). We test for spatial dependence, using several tests, such as Moran I statistics and Lagrange Multiplier test, as well as Robust Lagrange Multiplier test.

In chapter five, the multidimensional regional inequality in education is measured and analyzed. As in chapter four, for the case of health facilities and health status, chapter five also uses the method developed in chapter three to carry out the measurement and the analysis of welfare inequality for two components of education, namely education facilities and education enrolments. Unlike chapter four, however, chapter five does not explore the contiguity effects, nor the spatial dependence between education facilities and enrolments, because in most cases lower education is mandatory in Vietnam. Pupils are not able to seek for education in another school than the one they are by registration have to go to.

The measurement and analysis of regional inequality in per capita income is organized in chapter six, where the relationship between regional inequality and regional growth is explored and analyzed. A brief review of literature on regional income inequality and its link to economic growth is also given in this chapter. This chapter also provides an exploratory analysis of regional income inequality before presenting the method of measuring regional inequality in per capita GDP. As in the analysis of regional inequality in health and education, we use the one-stage decomposition method of Theil's index used by Akita (2003) to measure regional inequality in this chapter. Then the growth model, developed and used by Panizza (2002), is used to explore, analyze and detect the relationship between regional income inequality and regional growth in Vietnam.

Following chapter six, in which the relationship between regional income inequality and regional economic growth is considered, chapter seven is set to examine growth patterns of regional growth of per capita GDP in Vietnam, in order to see whether regional per capita income converge to each other or diverge from each other. In this chapter, the concept of convergence and divergence is elaborated on and discussed, and then the framework to test for conditional convergence developed by Sala-I-Martin (1995) is presented and extended for use. Also, a short review of literature on income convergence in general and the convergence in developing and transitional economies in particular is provided in this chapter. In addition, another test for convergence, using the unit root test developed and used by Evans (1998), is presented and used. In the last part of this chapter, the role of space is taken into account in order to investigate spatial dependence among provinces that could affect the regional growth rate in Vietnam.

Chapter eight is the last chapter of the thesis. It summarizes and discusses the main results and issues raised throughout this study. Next to the conclusions of the empirical findings, the main suggestion for further research and recommendations for policy purposes are discussed.

Chapter two
Macro Economic Reforms and Changes in Welfare in Vietnam

2.1. Introduction

Over the past 20 years, Vietnam's development has shown its mark in its economic history as the country has restored macroeconomic stability to achieve a high rate of economic growth and an impressive reduction of poverty (World Bank 2001a). These achievements were resulted from the Vietnamese government's efforts to recover the country from the deep socio-economic recession in the late 1970s and early 1980s. The recession has led the Vietnamese economy into hyperinflation and stagnation during the late 1970s and the first half of the 1980s that caused the deterioration in the people's living standard and widespread poverty. During the recession period, over three fourth of the Vietnamese population were living in poverty. Economic structure and infrastructure were badly damaged. Facing these difficulties, in the late 1980s the Vietnamese government announced a renovation program called *doi moi*[1] (the pro-market policies introduced at the 1986 VIth Congress of the ruling party) in order to restore the socio-economic conditions. Then, the *doi moi* has come into force after this congress during the second half of the 1980s. The main aim of the *doi moi* was to carry out the comprehensive socio-economic reforms. The reform has focused on transforming the economy from the central planned mechanism toward more market-oriented economic systems[2]. The pivotal point of the *doi moi* policy was macroeconomic structural adjustments, microeconomic reforms and transition to a market-oriented economy in order to foster growth and to improve people living standards and welfare. To achieve this transition, the government had implemented the comprehensive reforms from agricultural sector to industrial sector as well as gradually opened up the economy to the outside world. With a high rate of economic growth brought about by the reforms, people's welfare in Vietnam has changed. Poverty rate declined significantly as it is recognized that "almost no other country has recorded such a sharp decline in poverty in such a short period of time" (Government of Vietnam-Donor-NGO Poverty Working Group, 1999: iii), and household living standards have been considerably improved.

[1] The reform process taken place in Vietnam in the second half of 1980s (started in 1986) was known as renovation which is used in Vietnamese language as *Doi Moi*. Term *Doi Moi* has become very popular in the country and the world and is used as the official words in most documents.

[2] During the reform process, the key points in the policies for macroeconomic structural adjustments and microeconomic reforms are to change the structure of the economy in order to reduce its dependency on agriculture and import. Those were the policies toward constructing new economic zones, EPZs, and the equalization (privatization) of the State Own Enterprises (SOEs), etc.

As the data from the various household living standard surveys revealed the proportion of the population living below poverty line[3] measured by head count index significantly declined from 59% in 1993 to 37% in 1998, 29% in 2002 and merely 19% in 2004.

As the term "market-oriented economy" implies, the Vietnamese government has been trying to develop a well-functioning market economy to get rid of poverty (very high (75%) at the time of renovation announcement) and to catch up with other developing countries in the region and the world as well while maintaining the socialist ideology. Two salient dominant features of the economic model then in Vietnam as a market-oriented economy were the leadership of the Communist Party and the dominant role of the State-Owned Enterprises (SOEs) in its economy. This is a setback for the economic growth since most important enterprises such as post and telecommunication, civil aviation, electricity and petroleum, etc. were governed by hands of a small group of people who used their power to influence both the economy and politics.

In this chapter, the importance of the changes brought about by the reforms in the transitional period in the Vietnam economy and its impacts on macroeconomic performance that would have significant impacts on household's welfare will be discussed and elaborated. The chapter is structured as follows. Following this introduction, before elaborating the process of Vietnam economic development during slumping period in the second half of the 1970s and the first half of 1980s (before the d*oi moi*), we begin with an overview of the Vietnam geography in section 2.2, and the Vietnam demographic characteristics in section 2.3. These two sections are considered as rather important because they provide information for explaining patterns of spatial welfare distribution, which will be discussed later in other chapters in this thesis. The chapter is followed by the elaboration of the Vietnam's development process before the reforms (before 1986) in section 2.4. This period was characterized by the continuation of socialist transformation throughout the entire country after the North-South reunification in 1976 at the IVth Congress, which initially led to economic recession. The recession period had induced the government to implement the reform policy, which played an important role in the transitional process. This will be discussed in section 2.5. In addition, we will also discuss in details about the transition in which the reforms have been carried out in the agriculture (the sector that was well-known for the introduction of the "Product Contract System" or "household contracting"), and then in the industrial sector to which the "Three Plan System" and the Reform of the State-Owned Enterprises (SOEs) called equitization (an analogue to

[3] There are two poverty lines used in Vietnam namely the ***national poverty line*** and the ***international poverty line***. The ***international poverty line*** is derived by the Vietnam General Statistical Office (GSO) using the household expenditure survey data of 1993, 1998 and 2002 and has two components. The lower ***food poverty line*** is a measure of the expenditure per capita required to secure an intake of 2100 calories a day, regarded as the minimum nutritional requirement. Non-food items are added to obtain the ***general poverty line***. The basket of food and non-food items is determined by the consumption patterns of the first quintile of households in terms of capita expenditure. People whose expenditures are lower than the poverty line are considered poor. The 1993 total poverty line was VND 1.16 million per person per year. It was VND 1.79 million in 1998 and VND 1.92 in 2002. Meanwhile the ***national poverty line*** was developed by the Ministry of Labor, Invalids and Social Affairs. In the 1990s this line was the income equivalent of buying 15kg, 20kg and 25kg of rice per month as poverty line for food items. In 2001 the line was increased to include non-food items. Households are deemed poor if their income falls below the poverty line, which depends on where the household is located. The poverty line for urban areas is currently VND 150,000 per person per month while it is VND 100.000 per person per month for rural lowland areas and VND 80,000 per person per month for island and rural mountainous areas. The poverty rate is the proportion of households that fall beneath the line. In this study, we use the poverty line developed by the GSO of Vietnam called ***international poverty line***.

privatization) were applied. In the latter part of section 2.5, we will elaborate the impacts of the reforms on macroeconomic performance. In section 2.6 we will discuss about the changes that occurred in welfare brought about by the economic reforms. Finally, the chapter will be wrapped up in section 2.7 with a brief summary.

2.2. Overview of the geography of Vietnam

Vietnam is located in Southeast Asia, comprising the eastern boundary of the Indochinese Peninsula, with the natural areas of about 330,000 km^2. The country borders China to the North, Laos and Cambodia to the West, and the East Sea (South China Sea) to the East. Vietnam lies entirely within the tropical belt of the Northern Hemisphere, extending over 15^o of latitude. On the one end it is approximately 8^o from the equator, and on the other it is close to the Northern Tropic. The country stretches from its furthest North point at 23^o22N on the Dong Van Plateau, to its furthest South point on the Ca Mau cape, at 8^o30N. East and South Vietnam lay on the coastal line of the Eastern Sea, the Gulf of Bac Bo, and Thailand. The coastal line from Mong Cai at the border to China to Ha Tien is 3,260 km long (see the map shown in Appendix 2.2).

The country is at the intersection of several natural systems. It is greatly diversified in geology, terrain, climate, hydrology, soil types and fauna. Vietnam's territory is characterized by the crisscross of mountain ranges and high hills, presenting substantial obstacles to human access and thus constraining economic development. These mountain ranges and high hills are also characterized by low population density and high rate of poverty. However, the mountainous areas are often rich in minerals and hydroelectric potential as well as forest products.

The administrative map of Vietnam based on the 1999 census divided the country into 61 provinces (see appendix 2.1). These provinces are grouped into eight agro-ecological regions[4] namely the Red River Delta, the Northeast and the Northwest, the North Central Coast and the South Central Coast, the Central Highlands, the Southeast, and the Mekong River Delta (Appendix 2.1).

Through regions and provinces, the country has approximate 200,000 km of roads, of which over 15,000 km are classified as national highways, 17,000 km as provincial roads, and 36,000 km as district roads. Though the network is relatively well-developed, road quality is rather low for most parts of the country with many communes are lacking road access to all. The main highways of the country stretch over 1,700 km from north to south. The country's railway system is 2600 km long. It connects to the China's rail network in Lang Son in the Northeast and in Lao Cai in the North, and stretches on its main line through Hanoi to Ho Chi Minh City (hereafter is cited as HCMC). This railroad system is vital to the economy, carrying over 10 million passengers and transporting over 5 million tons of cargo annually (1999 data). In addition, there are over 12,000 km of waterways across the country, which play an important role in the cargo transport services to serve the economy. Air transport has been also developing and is becoming increasingly important to

[4] Initially Vietnam had been administratively divided into 7 regions for the purpose of surveys and poverty and inequality analysis. These regions are the Northern Mountains (including the North East and the North West), the Red River Delta, the North Central Coast, the South Central Coast, the Central Highlands, the Southeast and the Mekong River Delta. Latter, the country has been divided into 8 regions with the division of the Northern Mountains into two regions namely the North East and the North West as shown in Appendix 1 of this chapter.

the country. Beside the three international airports in Hanoi, HCMC and Da Nang, there are 13 domestic airports used for civil aviation. The domestic air services connect several of the generally less accessible areas to the main cities of the country.

The country surface is highly diverse and complex, changing from North to South and from West to East. Vietnam mountain ranges are regarded as a south-eastern extension of the Wen Nan and Qui Zhou plateaus of China. The prevailing orientation of many terrain features in the north and the central regions lies along a northwest – southeast trajectory. The western side of asymmetrical face of the Truong Son range slops down to the Mekong River valley, while the eastern side ends suddenly at the coast, causing severe floods and land erosion in the central areas of Vietnam. These kinds of terrain form several large plains in the country. The largest plain is the Mekong River Delta (hereafter is cited as MRD), which covers over 40,000 km^2. The second largest is the Red River Delta (hereafter is cited as RRD) with an area of around 15,000 km^2 (See map shown in Appendix 2.3). The prevailing terrain types in the mountainous regions in Vietnam are:

- The high mountain type: over 2,500 m height, concentrated in the northwest, partitioned terrain, slope of over 35^o, requiring strict protection measures with regard to sustainable watershed management.
- The medium mountain type: from 1,500 to 2,500 m height, partitioned terrain, medium slope of between 25^o and 35^o.
- The low mountain type: from 500 m to 1,500 m height, slope between 15^o and 20^o.
- The mountain plain type: rolling and hilly areas, with differences in elevation between 25 m and 100 m, generally favorable for forestry.
- The highland type (plateau form, popular in the central highlands of Vietnam): rolling and hilly areas, with differences in elevation of less than 25 m, generally highly favorable for forestry and agricultural production.
- The hill type: Absolute height of less than 500 m, relative differences in elevation between 25 m and 200 m, slope between 8^o and 15^o, generally favorable to forestry production.
- The peneplain type: Absolute height of 100-200 m and differences in elevation of less than 25 m, slope of less than 8^o, highly favorable to agro-forestry production.
- Carter terrain (limestone) types are not favorable for forestry or agriculture production.
- Mountainous valley and hollow types are generally favorable for agriculture production.

The mountainous areas make up about 75 percent of the total land area (see Appendix 2.3) as indicated by the dark color in the map). Vietnam is one of the most populous countries in the world, with more than 70 percent of its population (2004 data) living in rural and mountainous areas[5]. Most people living in such remote areas participate in a subsistence economy where products harvested from the forest serve as an important food source and provide materials for basic household items and household construction.

[5] Vietnam is ranked 13th in the world by population with over 84 million people in the 2006 estimate. By population density, Vietnam is at 47th place and is considered very densely because the country's geography is characterized by the large part of its area as highland and mountainous which are not suitable for agriculture. The population concentrated densely at the two main deltas as Hong river delta and Mekong river delta and other smaller plain lands throughout the country.

Approximately more than 20 million people (1999 data) live in or around forests and directly depend on forests for their livelihood. The forests provide a source of income through the harvest and sale of bamboo, firewood, herb medicine, fruit, fodder and game. However, they have become degraded and no longer present an abundant food source for the people living nearby. This reveals the close connection between the importance of protection of forest resources and poverty reduction in a largely rural population.

2.3. Demographic characteristics of Vietnam:

Information about the country's population can be gathered from various sources. In this study, however, the information is drawn mostly from the population census 1999 with proper adjustment made based on the VHLSS 2002 and VHLSS 2004. We structure this part as follows. The overview of the population density is provided in section 2.3.1 and the population characteristics in section 2.3.2. The section with education system and attainment is organized in section 2.3.3. The information about the household characteristics is given in section 2.3.4.

2.3.1. Population density

Vietnam's population is about over 86 million in 2007, rose from over 77 million as shown in the 1999 Census. The overall average national population density is not very high, at 232 persons per square kilometer (254 persons per km^2 in 2006). However, it is very unequally distributed across provinces and regions in the country due to the geographical characteristics. The provinces in RRD and MKD regions are particularly densely populated. Population density in the least densely populated provinces in the RRD region is still three times the national average. Hanoi is the one with the highest population density in the region - 2883 persons per km^2 (1999 Census). Generally, the highest population density is in Hanoi and HCMC. Commune/ward level densities of above 100,000 persons per km^2 are found in some parts of HCMC (presented as concentrated areas of darkish red/black in the map in the Appendix 2.4). In the MRD, Ca Mau is the only province with a population density lower than national average (215 people per km^2). All other provinces in the regions have densities lower than the national average.

The lowest population densities are found in Lai Chau (the North West), in the Central Highlands, and in a few areas in the South East (presented in yellow or light orange in the map at Appendix 2.3). Lai Chau, Son La, Hoa Binh, Ha Giang, Lao Cai, Cao Bang, Bac Kan, Lang Son, Kon Tum, Gia Lai, Dak Lak, Lam Dong, Binh Phuoc, Ninh Thuan and Binh Thuan are provinces with low population densities.

The reasons behind the unequal distribution of population are that RRD and MRD are areas with fertile and flat soil suitable for agriculture development. On the contrary, the upland provinces are mountainous areas, with rugged terrain, poor irrigation systems, and limited availability of cultivated land. Provinces in the Central Highlands have fertile soil. However accessibility to this region is not very easy. Therefore, this region remains sparsely populated despite the fact that the Central Highlands and the South East regions were the only regions in the country to attract a large number of immigrants during the end of the 970s, 1980s and 1990s due mainly to its availability of arable land in the Central Highlands, and the greater levels of industrialization and the high rate of urbanization and good market access in the South East.

For the provinces in the North Central Coast and the South Central Coast regions, people concentrate in the most-eastern areas along the coastal line, while the west is less densely populated since in the east the soil is more fertile and the land is largely flat. Proximity to the sea also increases the potential for the development of fisheries and salt production activities. The western part of these two regions borders the Truong Son mountain range, which is famous for the high terrain, the severe climate, the limited availability of cultivated lands and the poor soil quality.

In summary, Vietnam's population is very unequally distributed, and has the following main characteristics:

The population is largely concentrated in the two main areas of intense agriculture in the RRD and the MRD regions as well as along almost all the coastal areas, except that in Quang Ninh province, and between Phan Rang and Vung Tau towns, where density is lower than in other coastal areas.

The population distribution in Vietnam can be seen to gradually disperse as one moves from the low, coastal eastern areas to the rugged mountainous and partly forested terrain along the western and northern borders of the country. To a large extent, the country's relief is reflected in its population distribution with the large river valleys are the most densely populated, and less densely population distribution is found in mountainous and remote areas. The plain areas in the uplands, such as the highland plateau around Buon Ma Thuot are characterized by a distinctly higher population density.

2.3.2. Population characteristics

Dating back to 1980s and early 1990s, Vietnam is known as one of the least urbanized countries in the world with around 80% of its population living in rural areas. However, this figure has considerably changed over the last 20 years with the help of the reform progress. The ratio of urban/rural population is increasing. The share of rural population to total population declined from almost 80% in 1993 to 77% in 1998, 75% in 2002, and a further decline to just outside 70% (around 71%) in 2004. Among the eight regions in the country, the least urbanized are the Central Highlands, the North West and the North Central Coast. The most urbanized regions are the Southeast, which includes HCMC – the largest city in Vietnam – over half of the population in that region lives in urban areas (see Map in Appendix 2.4) and the RRD including Hanoi and Hai Phong.

Through its long history, the main religion practiced in Vietnam was Buddhism. However, it has been changing nowadays as about 60% of the population indicates no religious affiliation. Only close to 20% considers themselves as Buddhist. Christians make up about 9% of the population, while 2% belongs to other religious groups. The majority of the population (about 85%) is ethnically Vietnamese (Kinh). The second largest group is Chinese (about less than 3%). Others (including 51 other ethnic minorities) make up the remaining share of the total population.

According to the 1999 Census the proportion of the population aged five years old and below is 11.0 percent and that of population aged 65 and over is 5.8 percent. However, the geographical distribution of the population aged five years old and below is unequal. A comparatively low percentage of this age group can be found in RRD and MRD regions, and also to some extent along the coastal areas. On the contrary, a high percentage of this age group is found in mountainous and upland areas in the Northeast, the Northwest, partly along the Laos border in the North Central Coast region and particularly in the Central

Highlands, where the percentage of this age group accounts for one-fifth of the population (more than 20 percent) nearly double that of national average.

The research by Epprecht and Heinemann (2004) has shown that a large percentage of the population aged five years and below correlates directly to the areas with high birth rates. But to some extent it may also reflect a lower general life expectancy in areas with a larger proportion of younger people in the population, which coincidentally are also the regions with the highest poverty rates in the country. In addition, birth rates tend to be highest in areas where more ethnic minorities live. The study also found that the birth rates are highest in areas with the lowest population density.

The distribution group of people aged 65 years and over is also not equal among regions. The RRD, much of the coastal areas of the North Central Coast, and the northern part of the South Central Coast have a highest percentage of this group, accounting for 6-8 percent, and even higher than 8% in many areas. Some areas in the Northwest of HCMC and in the Northeastern part of the MRD regions also have proportions of this age group ranging from 6-8%. In contrast, in much of the Central Highlands, some areas in Binh Phuoc, Ninh Thuan and Binh Thuan provinces, in Ca Mau province of the MRD, and in Lao Cai and Lai Chau provinces of the Northeast and Northwest, there are less than 3% of the population aged 65 years and over – just less than half of the national average. This may be due to two factors as the Census revealed. First, it is the lower life expectancy among the population in these areas, and the second, there are, in some parts, the effects of high rates of immigrants into these areas.

The sex structure of the population in Vietnam is changing. The sex ratio of a country tells us the number of males in proportion to the number of females in a population. According to the 1999 Census, the sex ratio for whole Vietnam was 0.96. This means that for every 100 females in the population there were only 96 males. This ratio is lower than the international average probably due to more male than female deaths during the wars in the period 1940 - 1975 and 1979 - 1982 and from the high birth rate because of the people's preference for a son rather than a daughter. The natural sex ratio of the generations born after the wars has gradually been increasing from 0.94 in 1979 to 0.96 in 1999 (GSO 2002).

The sex ratio is also spatially distributed. Almost all the lowland and coastal areas have sex ratios less than one and those of most communes in the highland and mountainous parts of the country have a sex ratio above one. Exceptionally, the coastal areas are Quang Ninh province as a whole and most communes in Vung Tau and Phan Thiet, where the male population is more than the female, and much of the South East has the sex ratio above one. This unequal distribution of males and females among the provinces and the regions in Vietnam can partly be found in the patterns of labor migration. The provinces in the Central Highlands and Lam Dong, Quang Ninh, and Binh Phuoc provinces have a sex ratio above one, probably because of the following two reasons. Firstly, it is the impact of the immigration flow (among immigrating people, men almost always exceed women), and secondly, the high birth rates in recent years (the national sex ratios at birth is above one).

According to the 1999 Census, the percentage of literacy of the population aged 15 years and over stood at 90.3% nationwide. An urban-rural breakdown of the rate shows that 94.9% of the population in urban areas and 88.7% of the population in rural areas are literate[6]. This is a relatively high percentage compared with other countries in the region

[6] Literate population is defined in the Census as those who are able to read and write. We use the threshold of 15 years old because at this age, many people enrol in the secondary education, or enter the workforce; hence this can be considered a crucial point beyond which the ability to read and write is decisive in many ways.

particularly, and developing countries worldwide generally. A breakdown of the literacy rate by gender shows that the literacy rate of females is lower than the one of males. The literacy rate of the male and female population aged 15 years old and over was 93.9% and 86.9%, respectively.

Though the literacy rate is high, sharp regional differences can be found when looking at the spatial distribution of literacy throughout the country. As in the case of many other socio-economic indicators (elaborated above), the most visible contrast is between highland and lowland areas. Almost most coastal areas, all of the RRD, and most communes in the Southeast have rather high rates of literacy. The same picture is found in the MRD, in which the majority of the communes have a literacy rate of at least 95%. However, there are several clusters of communes in this region with rather low rates, with less than 50% of literacy can be found in these communes. Many parts of the Northeast have a literacy rate above 85%. The highest literacy rate can also be found in many areas of the RRD, and in some parts in the centre of the North Central Coast as in Ha Tinh province and in the southern part of Nghe An. As usual, most urban areas in the country have comparatively high rates of literacy among people aged 15 and over. In contrast to the high rate of literacy in the lowland areas, most of the upland areas have very low rates of literacy. In many places, the rate is even below 50%. The reason behind this is that in those areas a majority of the population belongs to one of the country's many ethnic minorities, which often do not have their own writing system, and certainly no formal education in any indigenous writing system while they face difficulties in getting fluency in the Vietnamese national language.

2.3.3. Education characteristics of the population

Before elaborating the education attainments in Vietnam, there is a need to have a brief overview of the current education system in the country. In Vietnam, children are able to enter the education system at a relatively young age. In many urban areas childcare is provided to children up to three years of age at crèches, and kindergartens are available for children aged between three and five. The five grades of primary education begin at age six. Students who finish grade five of the primary education go directly to lower secondary school (grades six to nine) without a need of passing a formal examination. To successfully complete grade nine, student must pass a formal examination. After completing lower secondary education (grade nine), students must also pass an entrance examination before they can enter the upper secondary level (grades 10-12). However, many students do not go on to pursuit the upper secondary education after completing lower secondary education. Instead, they participate in vocational training schools or professional schools. Those who pass the formal examination at the end of grade 12 earned an upper secondary diploma. These students could be able to sit in a formal entrance examination to enter higher education at universities and colleges.

The education level of the population is considered as an important determinant of economic growth (World Bank Vietnam 1993). Like the literacy rate, education attainment among population is different between age groups, sexes, and geographical regions. In this overview, the education attainment can be considered under several education levels prevailing in the country, such as primary education, secondary education, professional or technical educations, and higher education. However, the more focus is to put on primary and secondary education since these are the most popular in the country.

The percentage of people in a population that have completed primary school indicates the level of basic education that may be important for analyzing household welfare and development. According to 1999 Census, the percentage of the population aged 15 years old and over that has completed primary school was 42.5%. There is a difference in primary school level attained between the males and the females. The national average percentage of the female population in that age group that has completed primary school is 41.8%, and the national percentage of the male population that has completed primary school is 43.2%. However, the geographical distribution of the population that has completed primary school significantly varies. There are very clear contrasts between the northern part of the country – where the rates of primary completion are usually above 60% in the majority of the lowland communes, and in a larger number of communes in the less remote upland communes – and the southern part, where virtually all lowland commune have a rate of primary school completion even below 50%. While much of the uplands in the southern part of the country have a rate just below 20%, there is still a quite large number of communes in the southern part of the Central Highlands with a rate above 50%. The high rate of the primary school completion in the southern part of the Central Highlands could be the result of immigrants with higher levels of education, particularly the immigrants to these areas generally from northern Vietnam, where there was prevailing a higher level rate of basic education than that of local areas.

Also the findings from the 1999 Census show that the percentage of population aged 18 years and over that had completed lower secondary school accounted for 15.6%. For males in this age group, the rate was 18%, and for the females it was 13.4%. Compared with the education attainment at the primary level, these figures reveal that gender differences in education increase with the level of education. There are some patterns in the spatial distribution of lower secondary school graduation. Besides upland-lowland differences as usual in the primary education, the very interesting pattern found is the north-south discrepancy. The percentage of the population aged 18 years and over is higher in the lowland areas of the southern part of the country than in the north. This is particularly the case of the coastal areas, from Quang Tri province southward to the end of Khanh Hoa province and much of the lowland parts of the Southeast. The north-south differences are divided along the former Demilitarize Zone (DMZ), which divided Vietnam into two parts during the war against America[7]. As evidence from the Census, the areas with the highest lower secondary graduation rates are the lowland areas of South Vietnam prior to the reunification in 1975. On the contrary, almost all areas in the northern part of the country exhibit low percentages of the population aged 18 years and over that have completed lower secondary school – much lower than the national average. In some mountainous

[7] The DMZ was established as a dividing line between North and South Vietnam as a result of the First Indochina War against the French. Then, the Geneva Conference on July 21, 1954 recognized the 17th parallel as a "provisional military demarcation line" dividing the country into two states, Communist North Vietnam and Republic of South Vietnam. During the Second Indochina War (popularly known as the Vietnam War or War against US of the Vietnamese) it became important as a battleground between the North Vietnam forces on the one hand and the US and South Vietnamese forces on the other. Geographically, the DMZ spanned the whole width of present-day Vietnam (several hundred kilometres), and was a couple of kilometres wide. It reached across into a beach in the east. An island nearby was controlled by North Vietnamese forces during the Vietnam War.

provinces with a majority of ethnic minorities, such as Ha Giang, Son La, and Lai Chau, the rate is below 5%. A breakdown of lower secondary graduation rates by sex mentioned above, which says that there is lower rate for women than for men, appears to be true for all of the country. The rates are especially lower for females in areas such as in the MRD, in most parts of the Central Highlands, and in the mountainous parts of Nghe An, Thanh Hoa provinces. There is also the same story for north-south discrepancies between males and females in education attainment.

Moreover, the percentages mentioned above reflect general levels of education. What they imply is that a larger proportion of those who complete primary school continue their education, and complete secondary education, in the southern parts of the country than do primary school graduates in the north. Statistics show, however, that generally there is a higher rate of the secondary education completion among population in the northern part of the country, but there is a greater pool of people with a secondary education (as a percentage of the total population aged 18 years and over) in the southern part of the country.

Referring to the professional qualifications among population aged 15 years and over as the results from the 1999 Census revealed, there were merely less than 10% of the population in that age group that had some technical or professional skills. Put it in other words, there were more than 90% of the population that had never attended any courses of technical or professional schools. Among less than 10 percent of population with some technical or professional skills, however, the geographical and spatial differences can be found such as a urban-rural difference, and a north-south contrast. The areas with the highest rates of people with professional or technical qualification are found in and around the major urban areas, such as Hanoi and HCMC, also in nearby parts of most provincial towns, and industrial and export processing zones. This fact is not surprising since training opportunities for technical and other professional skills are mostly concentrated in urban areas. About the contrast between the north and the south, despite the fact (as revealed above) that there is a higher rate of the population with a completion of secondary education in the south than in the north, a larger percentage of northern population has some technical or professional qualifications than the southern population. A considerable portion of the population in the north received training in technical and professional skills at home or abroad (mostly in the former Soviet Union and Eastern European countries) in the communist period. This portion of the population makes the percentage of the group aged 15 years and over with some technical or professional skills increase in the northern areas compared to the southern part of the country.

In short, the education level in Vietnam is relatively high for such a poor country compared to other countries at the same level of economic development and even to other countries at the higher development level in the regions as well (Vietnam Human Development Report 2005, World Bank, 2005). This rather high level of education attainment may play an important role in explaining development patterns of the country. However, geographical differences in education attainments between sexes, age groups and ethnicities may also be useful for, and play an important role in explaining spatial poverty and welfare inequality prevailing in the country since the implementation of the market-oriented reforms. (See, among others, Vietnam Human Development Report 2001).

2.3.4. Household characteristics

Being considered as a unit for determining welfare, household characteristics play an important role in the welfare inequality and poverty analysis (World Bank Vietnam 2002). Therefore, a good and clear introduction to household characteristics will help in the process of determining and measuring poverty and inequality and their determinants, to be carried out in the next chapters of the study.

Household size in Vietnam has been changing. According to the VLSS 1993, average household size in Vietnam was 4.92. The size went down to 4.75 in 1998 (VLSS 1998) to 4.48 in 2002 (VHLSS 2002), and to 4.4 in 2004 (VHLLS 2004). However, the average household size varies across the country. On average, households are the largest in the North West, parts of the North East, in the mountains of Nghe An province, and in parts of the Central Highlands. Generally speaking, the household size is bigger in the mountainous areas than in the lower land areas. As the data from the 1999 Census and the VHLLS revealed, many areas in the RRD, parts of the coastal North Central Coast, and the South East have an average household size lower than the national average. Households with four or less people account for 55% of the total households in the country (VLHH98, GSO).

There is also an unbalanced trend between households headed by males and those headed by females. According to the data, the national average of households headed by females is 25.3%. However, this figure differs between urban and rural areas. The data show that the proportion of households headed by females in urban areas accounts for 39.1%, whilst that in rural areas is only 20.8%. The highest concentrations of female-headed households are in the southern parts of the country, particularly along the coastal areas of central Vietnam and the South Central Coast, and in and around HCMC (highest degree), in the South East and North-Eastern parts of the MRD. Particularly, urban areas such as Hanoi, Da Nang, and HCMC have the highest rate of female-headed households. The reason behind the fact of the high rate of households headed by females in rural areas is that most men often go to cities to earn for extra income apart from the income from agricultural activities so they left families to their wives to be heads of the households while for the households in urban areas headed by females as because with high education attained by women in the period of modernization and globalization, most women chose to raise a single family without married. This may imply that women might get better conditions through the high rates of urbanization in the country.

There are also large differences found in the education attainment between ethnic groups in the country. As the various rounds of the VLSS reveal, the education attainments in the Kinh people (Vietnamese people) is highest (around 95% at the national average) while that of the various ethnic groups (include Thai, Tay, Nung, Khmer, etc) except the Kinh is just above 50% in average (Vietnam Human Development Report, National Centre of Social Sciences and Humanities, 2001, 2004). The differences can be attributed to the cognitive competencies of people, and custom and tradition of the ethnic groups. Kinh people are considered to be superior to other ethnic groups in terms of attaining knowledge and adapting to situations and mostly reside in the lowland and more developed areas where the education infrastructure is higher developed (World Bank, 2004 and National Centre of Social Sciences and Humanities, 2001).

Referring to the education level attained by the heads of the household, particularly human capital in the form of education is considered as a fundamental determinant of economic growth (World Bank Vietnam 1993). Despite being such a poor country with a low level of income, general levels of education attained among its population are high, and the country also has a reasonably good level of education attained among household heads, and the levels are improving in the course of time. According to the 1993 data, there is only 12% of population that lives in households in which the head has no education, while another 38% lives in households in which the head has primary level of education. Put it in other words, almost half of population lives in households in which the head has at least secondary education or higher. These proportions have increased over times as shown in the results drawn from the household living standard surveys conducted in 1998, 2002 and 2004. The percentage of households in which the head is without any education marginally declines to 10%, 9% and also 9% in 1998, 2002 and 2004, respectively. Households in those the head has primary level of education also decreases from 38% in 1993 to 35% in 1998, 30% in 2002 and marginally down to 29% in 2004. In contrast, the population lives in households in which the head has at least secondary education or higher increases from 50% in 1993 to 55% in 1998, 61% in 2002 and marginally up to 62% in 2004. Across regions, the RRD has highest level of education attained by head, while the lowest levels are found in the remote Central Highlands. In General, levels of education of the head are lower in the South (South Central Coast, Central Highlands, Southeast, and Mekong Delta) than in the North (RRD and the North Central Coast). The education level by the head in mountainous areas is also lower than that in lowland areas and plain land.

In summary, as in the case of population education, the education attainment by the heads of household also varies between ethnic groups, provinces, regions, gender, and age, etc. However, as the various studies revealed, the large differences are between ethnic groups, between upland and lowland, between urban and rural areas. These findings may help explaining the development patterns between regions and between provinces in Vietnam.

2.4. A review of the economic development before *doi moi:* Period 1975-1985

After the war with the United States ended in 1975 and the reunification of the country in 1976, the economy of Vietnam had been characterized by the continuation of socialist transformation nation-wide, which started in 1958 in North Vietnam. This is marked by the Fourth National Congress of the Communist party in December 1976, which mapped out the period of transition to socialism throughout the country as “to carry out socialist industrialization and transform the economy from small scale production into large scale socialist production, and to give priority to development of heavy industry based on development of agriculture and light industry” (Fourth National Congress of the Communist Party in 1976, page 76). Vietnam’s economy was, in this period, performing poorly. In North Vietnam, two decades of war and isolation with inherent problems of applying soviet model central planning to a poor and subsistence agrarian economy had done little to improve the well-being of the population (VGOS 1992; Fforde and de Vylder, 1996). In the South, war and the distorted structure of incentives that had resulted from large and sustained volumes of US aid had impoverished many and enriched a few (Dacay,1986). Trapped in both low economic growth and dissimilar economic styles between the north and the south, the state quickly moved into the Second Five Year Plan[8]

[8] After the war with France (First Indochina War) ended in 1954, Vietnam had been partitioned into two regions: North and South Vietnam. In the recovering period after the war, the North Vietnam’s economy grew

(1976-1980) and the Third Five Year Plan (1981-1985) in order to apply the central planning upon the entire country (in the 1976-1980 plan), and to correct the shortcomings of the economy due to the application of central planning across the entire country in the Second Five Year Plan that caused a downturn in the late 1970s (in the 1981-1985 plan) as shown in Table 2.1. As the table revealed the decomposition of growth rate into different categories for 1979 and 1980 was below zero, the consumer price index (free market price) was so high, yearly-basic increase by more than 100%, and high levels of trade deficit. The main strategy of the Fourth Congress of the Communist Party was the continuation of the socialist transformation nation-wide (applied to the south), but the pace of the process would be slowed down. The two main goals set by the Congress were: i) to build the system of collective command; and ii) to construct production at a large scale.

The Second Five-Year Plan, from 1976 to 1980, contained many points of voluntarism such as: i, to strive for achieving rapid development of agriculture in order to provide food to its people, to improve people's living standard, and to accumulate capital for industrialization production (agro-based developing model); ii, to develop heavy industry to support agriculture and then to increase capital accumulation; and iii, to implement basic socialist transformation to the south. In this period, the economic model developed in the north was applied to the south as the government nationalized state enterprises and cracked down private businesses, organized farmers into the northern style of agricultural collectives.

In the Third Five Year Plan from 1981 to 1985, in order to reverse the bad situation of the economy at the end of the previous plan, the country attempted to break the rigidity of the centrally planned mechanism. This attempt had been accepted by the country top leaders. Thus, a new and relatively liberal resolution had been introduced as to encourage the development of a private and household economy. Most important breakthrough was the introduction of the product contract system in agriculture, a step further away of the agrarian reform which had been put forward in the late 1979s by passing Degree No. 100. In the product contract system, households, rather than cooperatives were considered as the main economic unit, and were assigned lands in a cooperative by signing a contract to deliver a given amount of output at fixed prices as quota for using the cooperative land. The households were allowed to retain and trade surplus output beyond their quota. The cooperative remained responsible for providing various services as ploughing, irrigation and marketing, and supply of seeds, fertilizers and pesticides.

In the industrial sector, a significant change was the application of market-oriented reform, whereby the state-owned enterprises (SOEs) were allowed to operate under three plans in the Three Plan System Mechanism. Under Plan One, the enterprises were provided with the inputs at subsidized prices, but were required to supply set quantities of goods to the state. Under Plan Two, the enterprises could produce beyond the amount specified in Plan One and were able to buy additional inputs needed. Plan Three allowed enterprises to engage in sideline activities as they were permitted to produce and sell surplus products from freely purchased inputs.

---

well, quickly recovering agriculture and transportation. Since 1958, the North has realised socialist transformation in the economy, establishing centralised-economy developing model and nationalising all means of production, planning the national economy centrally. In this period, the North government implemented the First Five-Year Plan which took place from 1961 to 1965, the country moved into a constructing period under central planning.

However, most of the targets in the second five year plan were hard to meet. Collective mechanism in agriculture and nationalization of industry proved its shortcomings and there were many setbacks as the growth of agricultural and industrial sectors in the second plan was very little. However, with the introduction of the Product Contract System in agriculture and the Three Plan System in industry in early 1980s – the beginning years of the Third Five Year Plan – the economy recovered and was in much better prospect. As table 2.1 indicates, total social product and industrial outputs growth rates were negative during 1980-1981, then increased in early years of the third plan before falling again in 1985-1986. For the period 1976-1980, annual growth of agricultural output was around 1.9 percent, significantly lower than annual population growth which stood at 2.3 percent (VGSO 1991). However, total agricultural output increased in 1981, 1982 before declining again during 1984-1986. There was no different story in the industrial and trading sectors. Nationalization of industries and commerce was not successful. The industrial sector had a very low rate of growth. Industrial production was stagnant with an annual growth rate of about 0.4 percent only (VGSO 1991). As a result of mismanagement and misleading policies during the time, the country's economy was in very bad condition. Food shortage was so critical that it forced the government to import 1,576 million tons of food grains in 1980 despite being an agricultural-based economy. Budget deficit was very high. During the period between 1976 and 1980, the government deficit was 2.5%. This figure rose to 14% in the period 1981-1985. The peak of the deficit was in 1985 when it reached 36%. Gross investment was too low. Prices rose almost more than 50 percent annually and import was much higher than export. The country was in short of investment capital for the economy leaving many plans unfinished and leading to serious shortage of essential consumer goods. Table 2.1 indicated, after a slight recovery in the early years of the 1980s, the economy was in recession again in the mid 1980s as the selected macroeconomic indicators in the table showed.

Table 2.1. Selected Macroeconomic Indicators, Period 1976-1986

Indicators	76-80	81-85	1979	1980	1981	1982	1983	1984	1985	1986
Growth rate (% per year)										
Total social product[a]	0.5	6.4	-2.0	-1.4	2.05	8.9	6.7	8.3	5.7	2.2
Per capita social product	-1.8	4.2	-4.2	-3.6	0.3	6.7	4.5	6.1	3.5	2.1
Total industrial output	0.6	9.5	-5.5	-1.4	1.0	8.1	12.8	13.1	11.9	6.5
Total agriculture output	1.9	4.9	1.7	5.2	4.9	10.9	7.0	4.2	4.7	0.3
Retail price index	60.0	74.2	119.4	125.2	69.6	95.4	49.5	64.9	91.6	590.0
Gross investment (as % of net material product)	13.1	13.0	14.0	13.3	11.7	10.8	12.8	14.9	15.0	8.0
Budget deficit (as % of total govern. expenditure)	2.5	14.0	5.2	1.0	17.5	8.0	4.5	4.0	36.0	30.0
Trade deficit (% export/import)	24.2	33.8	33.1	30.3	29.1	32.6	37.5	35.0	35.0	37.0
Per capita staples production (kgs)	259	295	266	268	273	300	296	303	304	301
Unemployment rate (%)	12	20	13	15	29	32	25	19	10	9

Notes: a. The net material product national accounting system is explained in endnote 111.
Sources: GSO of Vietnam and the World Bank Vietnam: various years

During the time prior to the Sixth Congress of the Communist Party in 1986, the Vietnamese Government was facing two difficulties. First, the process of the transformation of the economy was in stagnation. In the beginning period, especially in 1979 and 1980 there was no output growth but the population grew very fast. Real capita income continued falling from an already low level in the years of 1979, 1980. Per capita total product declined 2.0% and 1.4% for the 1979s and 1980s respectively (See Table 2.1). Second, the partial and gradual market-oriented reform was seriously defected. There was output response but the macroeconomic imbalance which led to hyperinflation undermined the support of the reforms. In the years 1985 and 1986 the price rose almost from 100% to 300 % in the free market retails. In order to cope with the downturn, in 1985 the state introduced a comprehensive reform in currency, price and wage. The number of goods subject to price controls was reduced, and efforts were made to eliminate subsidies to producers and consumers. However, the efforts to reform subsidies and prices resulted in heavy losses for SOEs which were covered by the state. The coverage of SOEs losses had a dual impact. First, it boosted the budget deficit, which rose to an unsustainable 12% of GDP in 1985 (World Bank, 1990: Table 2.1). Second, the budget constraint of the SOEs stimulated further inflation. The GDP deflator, which rose from 307 in 1984 to 588 in 1985, took off to reach 3415 in 1986 (World Bank, 1990; Table 2.1). As a result, hyperinflation prevailed and domestic saving collapsed (Fforde and de Vlyder, 1996). The economy was again facing severe stagnancy and downturn. Therefore, in December 1986, the Vietnamese Government decided to change the course of the economy reform toward more intensive reforms to transform the economy from a centrally planned to a market-oriented economy. This transformation had been known as renovation or *doi moi* in Vietnamese. The *doi moi* suggests not a full-scale conversion to capitalism, but rather a cautious acceptance of the market as a means for achieving economic growth, improving people's living standard and, thereby, maintaining and strengthening the party's political and economic control. The main goal is to gradually transform the economy to a socialist market-oriented economy through step-by-step procedures.

2.5. Economic reform under *doi moi* and its impacts on macroeconomic performance

In 1986, the economy was in recession again. The severe macroeconomic imbalances of the mid-1980s exposed the inherent contradiction of a state-led, market-subordinated development strategy (Sepehri and Akram-Lodhi, 2002). In addition, the advent of hyperinflation led to a collapse of real spending and a liquidation of domestic saving. Whilst the state's earlier attempts to sustain central planning and its institutions succeeded in slowing down the growth of the non-planned economy, the symbiotic relationship between the planned and non-planned sectors created more pressure for further intensive reform. Therefore, in 1986 the Vietnam Communist Party launched a comprehensive reform program called renovation or *doi moi* in the entire economy. The fundamental viewpoints of the *doi moi* were: i) to develop a multi-sectoral economy; ii) to shift the economy from a planned centralized and state subsidized mechanism to a socialist-oriented market economy under management of the state; iii) to enhance foreign economic efficiency on the basis of expanding economic cooperation to the countries in the region and throughout the world; and iv) to democratize all fields of the socio-economic life, to implement administration reforms, to improve people's living standard and building a good welfare system for the people.

This section is structured as follow. An overview of a comprehensive economic reform will be provided in section 2.5.1. In this part, the process of the reform in agriculture, industry, foreign trade and pricing will be examined. In section 2.5.2 the impacts of the reforms on the macroeconomic performance will be analyzed. Reforms will be elaborated for various periods according to the different phases of the process of economic reforms in the country.

2.5.1. Economic reform under *doi moi*

The comprehensive reforms toward market-orientation in the Vietnam economy had been launched by the Vietnam Communist Party in 1986 after the economy experienced the unprecedented downturns brought about by the earlier reforms in the form of socialist transformation in the second half of the 970s and early years of the 1980s. This time, the reforms had been placed on agriculture, industry, foreign trade and pricing. The key reforms were microeconomic. The most visible set of reforms began in 1987, when price controls imposed on all commodities such as rice, kerosene and some public utilities were lifted. In the same year, internal trade was significantly eased when control posts were abolished, providing a boost to international trade. The key reforms on foreign trade were embarked in 1988, when tariffs began to replace quantitative restrictions (Nghia *et al.* 1999). Concurrently, the government ceased its exclusive control of foreign trade through state trading companies and import export licenses.

The start of the reform was in agriculture by applying a set of microeconomic reforms to this sector in the late 1980s. The most important measure was the introduction of Degree 10 in 1988, which formally de-collectivized agriculture. The reform further went beyond the contract system which had been set up in early 1980s. The quota system, in which cooperative land was allocated to households which signed a contract to deliver a certain set amount of outputs at fixed prices, on household production was eliminated, allowing farmers to make all decisions regarding resource allocation, production and sales. Crop and input prices were liberalized. Land tenure had been given to the farmers in 1988. Next to the land reform, the new land law was passed by in 1993, which classified those farmers who had the right to use land distributed to them for 20 years and that this right could be renewed. This gave more opportunities to farmers to sell or mortgage the rights to use their land (World Bank 1993b, chapter 2). The agricultural sector, which was the dominant sector of the economy at that time, was effectively privatized.

In the industry sector, there were only few reforms initially. Because of the so-called socialist orientation of the market economy applied to Vietnam, there was no attempt to reform and privatize the state enterprises. However, private economic activities did bloom in the informal commercial and service sector. SOEs had been granted more autonomy by primarily removing the role of planning targets in the decision making processes of these enterprises. Of greater importance in the reforms in this period were the efforts to strengthen the economic role of the private sector. Prior to the *doi moi*, the private sector had not played a significant role in the transition process. Another important step was the elimination of the state monopoly of foreign trade in 1988, allowing the establishment of Foreign Trade Organizations (FTOs) and permitting some firms to engage directly in international trade outside the FTOs. The economy had been opened up to foreign direct investment by the introduction of the first law on foreign investment in 1987.

There were strong economic responses to the market-oriented reforms. Once again as in the early 1980s as discussed above, macroeconomic imbalance continued to undermine the economy. During the period, from 1986 to 1989, inflation was extremely high. The inflation rate in 1986, 1987 and 1988 was 487 percent, 301 percent, and 394 percent respectively. The cause of the macroeconomic imbalance was perhaps the deficits of the public sector. The rapid expansion of credit resulted in a sharp increase in prices. The higher the rate of inflation was, the more people shifted from Vietnamese dong, the local currency, to use dollars and gold, which was circulating freely at the time. This behavior had triggered further instability of the economy. The inflation also undermined international competitiveness, with the dong significantly overvalued in real terms. The trade balance was also increasingly in deficit, up to about 10 percent of GDP in 1989 (Ministry of Trade, 1990). Although trade with the non-socialist countries covered only 26.8% of import bills over this period as compared to 31% over the period between 1981-1985 (Phong and Beresford, 2000). Hence, three years after the beginning of the reform under the *doi moi*, the economy was, once again, facing a recession.

Under the pressure of an impending crisis, the government decided to accelerate the process of transition to a market-oriented economy with a combination of structural reforms and stabilization measures. The main structural reform, adopted in early 1989, included the elimination of price control and the system of state procurement. The dual pricing system under the three-plan system in the course of industrial reform destroyed both efficiency and stability. Firms were to sell at low prices and they then had to be financed by credit from the central bank. This led to the creation of money and, as a result, inflation.

Facing extremely high rates of inflation, the country adopted the stabilization program in 1989. Two key components of the program were to raise and to stabilize interest rates, and devaluate and unify the exchange rate[9]. In this period, the central planning was abandoned. The authorities readily implemented an IMF type stabilization program without IFM funding. The program called, among other things, for a tight monetary policy, a reduction in government spending, deeper, more profound market liberalization, the introduction of user fee for publicly provided services, the broadening of the tax system, and a drastic devaluation of the dong (Ljunggen 1993: Fforde and van Vylder 1996). As a result, inflation was almost brought to a halt by mid 1989. Credit continued to grow but there was a substantial portfolio shift from dollars and gold to dong assets. These considerably helped to stabilize the economy.

In parallel with these measures, in order to reduce the government deficits perhaps caused by the growth of the public sector spending, the government chose to raise taxes and to reform the public sector. In fact, the tax base in Vietnam was very weak. Most of government revenue was from taxing the state enterprises, and the price reform undermined the profitability of these enterprises. The only way that the government was able to bring the budget deficit down was to cut government spending in 1991 and to cut short the credit subsidized to the SOEs sector, to improve the fiscal burden and budget deficits. In this process, the government launched the SOE reforms program called equitization in 1989. Before *doi moi*, the country had over 12,000 SOEs. Economic performance of these

[9] Vietnam had a dual exchange rate system prior to the intensive reform in 1989: the official exchange rate and the swap rate. The official exchange rate referred to the exchange rates between Vietnamese Dong and other foreign currencies, as set by the State Bank of Vietnam. Whilst, the swap rates were the rate applied for the purpose of facilitating purchase and sale of foreign currencies between two business entities at non official rates. These swap rates were determined by market supply and demand.

enterprises lagged as they lacked incentives to be more efficient and profitable. Accumulated bad debts of poor performing SOEs led to increased fiscal burdens and budget deficits that caused a protracted problem[10] to the Vietnamese government, as the devaluation of the dong, sky-rocketed inflation.

In the process of SOEs restructuring, there were about 800,000 workers – one third of state enterprises labor forces - left the sector, and the number of firms considerably declined to 7,000 in 1992 and to around 4.800 in 2003 from 12,000 SOEs in 1989. These policies gradually brought the expansion of credits under control. In two years, 1990 and 1991, they reduced the deficits by six percentage points of GDP by cutting subsidies to state enterprises, reducing investment programs, constraining wage increases to below inflation, and also demobilizing one-half million soldiers. As a result, money growth was under control, and inflation was cut down to a single-digit level. In 1995 domestic credit increased by 20.3 percent, none of which went toward the budget, credit to state firms increased by 16.7 percent, and credit to the private sector increased by 37.2 percent. All these policies along with restrained monetary policy succeeded in bringing inflation down to a one digit figure.

Another important element in Vietnam's successful stabilization program was the growth of exports. Vietnam was critically dependent on some key imported inputs like steel and fertilizer, which came mainly from the Soviet Union and other socialist countries in the Eastern Europe. However, the fall of the Soviet Block has negative impacts on Vietnam's import of these inputs. Without foreign exchange, it was difficult to import inputs to serve its production. Therefore, shortage of foreign exchange for import propelled the government to strive for a proper solution. The success of Vietnam in continuing the import of inputs was achieved through the liberalization of trade and devaluation of the currency to stimulate the growth of export. As a fact, Vietnam was known as one of the transition economies most open to foreign trade and investment (Desai, 1998).

The government policies toward reforms both in agriculture and industry, and trade liberalization brought successes to the economy. The liberalization in agriculture met with positive responses, and within one year, has transformed the country from a rice importer to one of the largest rice exporters in the world. Rice output increased around 25 percent per year that played an important role and had a very positive impact on exports to convertible currency areas. Furthermore, as also a crude oil exporter, when oil became the most rapidly growing export in 1992, this contributed substantially to the country's export growth. Nevertheless, industrial sector still played a very small role in the export, accounting for just 13 percent of total export in 1992 (Ministry of Trade, 1993). Increases in export of these products partly help the country to solve the problem of foreign exchange, and to partly meet the demand import of inputs. With the effectiveness of the very first foreign investment law, the FDI was blooming in Vietnam during 1990s (see Table 2.2).

[10] In his speech at the National Conference for accelerating SOE reform on March 14 2004, then PM. Phan Van Khai Said: "Not many SOEs are profitable. The taxes collected from SOEs are just VND 8 billion while the state budget is VND 87 billion. SOEs' bad debts are high as 8.5 % while the average rate for the economy is just 6.1%. Total receivable and payable debts of SOEs total VND 300,000 billion (around US$ 20 billion, more than 50% of Vietnam's GDP). I have had to settle bad debts of SOEs at least twice since I was Deputy PM. After just settling VND 18,000 billion, I found another VND 18,000-19,000 billion bad debt "returning". If this situation continues, the growth rate of the economy will be affected."

Table 2.2. Selected Macroeconomic Indicators, 1986-1993

Indicators	1986-1988	1989-1993
	Growth rate (% per annum)	
Real GDP	4.4	6.5
Per capita real GDP	2.0	4.8
Inflation (CPI)	365.3	38.9
Unemployment rate	14.6	12.7
	As a share of GDP	
Gross investment	8.4	16.4
National saving	-1.4	11.3
Government deficit	5.9	4.9
Current account balance	-9.8	-5.1
Per capita staples production	296	332

Sources: GSO of Vietnam and the World Bank Vietnam: various years

2.5.2. The consequences of economic reforms on macroeconomic performance

The reform process under the *doi moi* brought about tremendous successes to the Vietnam economy. Certainly, its achievements exceeded what the reformers imagined when they launched the *doi moi* program. To go through the achievements brought by the reform process, we will examine its outcomes in different stages of the development process.

2.5.2.1. Economic achievements during the 1986-1993 period

Despite the severe recession in the early years of the second half of 1980s, the Vietnamese economy showed the sign of recovery in the late 1980s. Through many solutions and decisions issued by the government in an attempt to improve economic management, monetary and agricultural policies, the economy started to grow. During the early years of this period, however, the fact that the old mechanism did not yet disappear and the new one did not prominently emerge, made reform more effective. Despite these, the economy started to show a sign of recovery. As Table 2.2 reveals, on average, GDP increased by 3.9 percent per year during these five years, per capita GDP increased 2% during 1986-1988, and 4.8% during 1989-1993 despite skyrocket increase consumer price, which rose to 365.3% per annum during 1986-1988.

Since the late 1980s, there have been evident changes, especially some positive changes in agriculture. As mentioned above, in 1988 the government put forward the new regulation not to fix farm output quotas to each household, and to consider each household as an economic unit. This new regulation brought positive changes to the economy. In the food grain production, instead of having to import 450,000 tonnes of food as the country did in 1988 and before, Vietnam became one of the rice exporting countries one year later in 1989 - the year that the country exported nearly one million tonnes of rice - and became the world third biggest rice exporter with export volume reaching 1.5 million tonnes in only one year 1990. In the industrial sector, key industries such as electricity, laminated steel, cement and crude oil attained fairly good growth.

During this plan, average import-export value decreased by 28 percent, gradually reducing the trade deficit. The ratio of export to import during the 1986-1990 period was 1 to 1.8 in comparison with 1 to 4 in the period 1976-1980. Another great success was that inflation had been kept under control and driven back from 774.7 percent in 1986 to 223.1 percent; 34.7 percent and 67.4 percent in 1987, 1989 and 1990 respectively.

In short, the successes of the renovation in the 1986-1990 periods were a recovery of production, a growing economy and rolled-back inflation. More importantly, there was a fundamental shift to a new management mechanism as application of market-oriented mechanism and private ownership, implementation of trade liberalization and multi-sectoral economy. Foreign direct investment (FDI) started to accrue into the country to help boosting production, generating income and employment, and eventually improving people's living standard. These successes make the process more important as the reform had been implemented before Eastern European countries and the former Soviet Union got into total recession. However, it took the country almost 5 years more to get rid of the social-economic crisis which broke out in the first half of 1980s.

2.5.2.2. From 1994: period of high economic growth

With the initial achievements attained during the period from 1986 to 1993 as mentioned above, the confidence of the country had been boosted. The government continued to work out the strategy toward stabilizing and developing the socio- economic conditions, putting forward the orientation and tasks for the period thereafter. The major setbacks the economy faced during end of 1980s and early 1990s were the long lasting economic embargo together with economic blockade imposed by the US while the Eastern European countries were facing serious crises during their transitional period. All these factors had a negative impact on the Vietnamese economy. The total foreign trade turnover to these countries sharply decreased (by almost six seventh). The two-way trade turnover in 1991 was accounting for only 15.1 percent of that in 1990. However, it was of great advantage that the renovation started to have effectiveness; economic units were gradually adapted to the new management mechanism. There were great achievements during this period brought about by the reforms as follows:

1. Fundamental changes in the economic mechanism. There have been many sectors functioning in the economy namely state-run, state capital, private capital, cooperative, individual, etc. in which the non state-run sector accounted for 60 percent of GDP. Economic sectors were handed with rights of land use and export-import activities. The state-run sector, however, was still given special attention to help it play the decisive role in the economy.
2. Economic structure reform: In pursuing the high rate of economic growth, the country continued the course of reform in its economic structure. To do this, the government paid attention to raise the proportion of service and industry and to production, steadily reduce that of fishing, forestry, and agriculture. Economic structure began to shift towards establishing essential areas, central industrial zones, and export processing zones as well as areas specialized in industrial plants and producing food and beverage.
3. Inflation was kept under control and driven back. With the development of production, convenient commodities circulation and anti-inflation experiences from

several years before, prices were gradually stabilized. Prices of goods and services increased by 67 percent in 1991, 5.3 percent in 1993, and only 4.5 percent in 1996.

4. Establishing economic-commercial relations with more than 120 countries that led to increase foreign trade turnover by more than 20 percent a year. Non-refund aids and loans for social-economic development investment by many countries and international organizations have been granted. The total ODA loan that Vietnam received from 1994 to 1997 was 8.53 billion dollars. In the field of foreign direct investment, during the time span from 1988 to the end of 2005, there were more than 4,000 licensed investment projects with total capital of approximately 50 billion US dollars (Ministry of Planning and Investment, 2005).

The achievements resulting from the changes in economic structure and mechanism had created stability and fundamental changes in the economy. As a result, fair and high economic growth rate has been attained for the last 15 years, unlike other transitional economies in Eastern Europe and the former Soviet Union where economies experienced a dramatic decline in the initial period of the transition (Ivaschenko, 2001). Vietnam was considered as a success story with high rates of economic growth without any decline (even from the very beginning of the transition), macroeconomic stability, stable prices, and increasing annual rate of employment over the last 15 years despite a marginal slump in a very short time in 1986-1987 when GDP growth reduced to 0.7 percent from the level of 2.3 percent in one year earlier[11]. Since 1990, GDP growth rate in the country was always sustained at high level. During the first half of 1990s, the average GDP growth rate was kept at 8.2 percent annually. Thereafter, since 1995 annual economic growth in Vietnam was still kept at high level. In addition, the economy was also registered a very impressive annual growth rate of employment and a low level of inflation rate. As Table 2.3 shows, between 1996 and 2004, the rate of economic growth was always at more than 7 percent on average, and employment rate increased during the period by about 2 to 5 percent. The growth performance of Vietnam has been among the best of transitional economies both in Eastern Europe and in Asia. Real GDP in 2001 was approximate 50 percent higher than in 1995, and real GDP in 2004 was also 50 percent higher than in 2001. It means real GDP has doubled after 10 years from 1995 to 2005. Meanwhile, as shown in Table 2.3, year-on-year changes in consumer prices kept at single digit, and stable and low inflation rate for over last 10 years indicate stability in domestic market.

These successes of *doi moi* in Vietnam during the last two decades of reforms have changed people's welfare. Whilst the Eastern European countries faced a downturn trend in economic growth in the beginning years of transition, Vietnam's economy did not suffer from this tendency. The Vietnamese economy had sustained a fair growth rate from the beginning years before attaining a very high rate of growth for over the last 10 years. We cannot deny the fact that there were also negative impacts on the people's welfare created by the reforms in the beginning as people lost their job during the SOEs reforms, which negatively affected the people's income and household welfare. However, the benefits

[11] Before 1988, Vietnam's national accounts were calculated using the Net Material Product System (MPS). Subsequently, Vietnam followed the System of National Accounts (SNA) used by most countries and has calculated GDP in 1989 market prices for 1988. The MPS data for the pre-1988 period excluded some services counted as part of GDP in SNA. More importantly, production data were aggregated using the rather arbitrary fixed plan prices of 1982.

gained from the reforms have surpassed the loss. As shown in Table 2.3, the GDP per capita increased almost double fold over the last 10 years which led to a sharp fall in the poverty rate over times. According to VLSS surveys conducted in 1993, 1998, 2002 and 2004, the poverty rate in these years was 59%, 37%, 29% and 19%, respectively.

Table 2.3. Selected macroeconomic indicators: 1996-2006

Indicators	1996	1998	1999	2000	2002	2003	2004	2005	2006
Real GDP growth rate (%)	9.3	5.7	4.8	6.8	7.0	7.1	7.5	8.4	8.2
Changes in consumer prices (percentage change)	n.a	9.2	4.0	7.0	3.8	4.0	9.0	8.4	8.0
Employment (year on year changes in percent)	2.21	2.14	2.11	2.01	3.5	4.2	5.0	5.3	5.4
Inflation rate (y-y basics)	4.5	9.2	0.7	-0.5	2.9	3.0	2.7	8.9	7.8
Unemployment rate (%)	5.88	6.85	6.74	6.42	6.01	5.78	5.60	5.31	4.82

Note: n.a indicates data are not available.
Sources: General Statistical Office, various years

2.6. Changes in people's welfare

Vigorous economic growth for over a long period as the results of a comprehensive economic reform called *doi moi* or renovation since the beginning of the *doi moi* in the mid 1980s had brought changes to living conditions and people's welfare in Vietnam. Over the last two decades, per capita income had increased almost four times. This high growth rate has considerably reduced poverty and, at the same time, brought changes to the state of welfare distribution in the country. It would not be surprising, given such a good economic performance, to see that many economic and social indicators have improved considerably during the period. Table 2.4 illustrates the extent of the improvement in a range of welfare indicators. As the statistics provided in Table 2.4 reveal, all the key welfare indicators have been significantly improved. The second row of the table shows the real GDP per capita consumption has increased close to four times between 1992-93 and 2004-05. The poverty rate has also declined considerably, from 58% in 1992-93 to only 19.5% in 2004-05. Other hey welfare indicators representing education, health and basic needs have also improved that would lead to a better welfare for people in the country.

Table 2.4. Key welfare indicators during doi moi

Indicators		1992/93	1997/98	2004/05
GDP per capita consumption (real) (USD)		180	311	720
Poverty rate		58.1	37.4	19.5
Lower secondary enrolment rate (%)		30	62	90
Doctors per thousand population (doctor)		26	38	50
Child malnutrition (%)		50	34	26
Adult malnutrition (%)		32	27	19
Access to clean water (%)	Rural	17	29	40
	Urban	60	75	81

Source: General Statistics Office of Vietnam

Consequently, the improvement in welfare indicators have raised questions to those concerned about how the patterns of welfare distribution during the years of economic reforms, and how welfare inequality have evolved during the *doi moi*. Previous studies on household expenditure distribution using Gini index of inequality have found the increasing trend in household expenditure inequality (Dollar et al., 1998; Binh Nguyen et al., 2003; World Bank and the GSO, various years) using the data from various household surveys. Initial estimates of the Gini coefficients for income and expenditure inequality representation for different income categories (Table 2.5) show the increasing trends of income inequality in Vietnam. In Table 2.5, we computed the Gini coefficient for various types of income and expenditure in order to represent a picture that shows the pattern of income and consumption of the people in the country. These breakdown calculations would further help people understanding the way Vietnamese people spend on their food, non food items. As the breakdown Gini has shown, inequality in total income increases over the times but it is more moderate than the inequality in non-wage income but inequality in non-farm income has declined between the 2002 and 2004. In the case of the Gini coefficient for expenditures, as statistics reveal, inequality in total expenditure has increased over the times but that of non-food expenditure and durable consumption declined from 1992 to 2004 but Gini for non-durable consumption has increased. These findings reflect the consumption patterns of people in Vietnam as the country records a faster growth, economic situation is improved. As a result, people's income increases which leads to increase in their living standard. It is in line with the consumption behavior theory as that people consumption pattern changes towards spending more on luxury good when they become rich.

Table 2.5. Gini coefficients for overall inequality in income and expenditures

Gini coefficients	1998	2002	2004
Total income	-	0.409	0.414
Income excluding wage	-	0.496	0.498
Non-farm income	-	0.609	0.570
Total expenditure	0.351	0.368	0.372
Non-food expenditure	0.505	0.481	0.472
Durable consumption	0.634	0.596	0.564
Non-durable consumption	0.336	0.351	0.360

Source: Own calculation, based on data from VLSS 97/98, VHLSS 2002 and VHLSS 2004

2.7. Summary and conclusion

In chapter two we provided an overview of the economic development process in Vietnam since the reunification of the country following the end of the war against the US in 1975. In 1976 the country started to apply the soviet model already existing in the North to the South. The main economic mechanism was centrally planning. However, this economic model performed poorly. The Vietnamese economy was in a severe downturn at the end of the 1970s and the first half of the 1980s. Economic growth attained during this period was very low, and even below zero. Inflation was very high and protracted. Trade and government deficits were high. All major macroeconomic indicators indicated a stagnant economy. The agricultural and industrial production was low and led to food shortages for many years in the 1980s despite the agrarian character of Vietnam. The entire economy was in an impasse.

In the midst of the recession, the government of Vietnam launched a comprehensive reform program in 1986 to strive for overcoming the economic crisis and stabilizing the macroeconomic condition. Initially, the reforms were a step to partly abandon the central planning. However, because the old mechanism did not completely give way to the market –oriented mechanism, the economy still faced a downturn in the second half of the 1980s despite the initial recovery recorded. However, the *doi moi* succeeded in getting the process of institution building and marketization on track. The government, once again, took a step further to push up to the more intensive reform in the early 1990s. The intensive reforms had been carried out in agriculture, industry, foreign trade and pricing. The foreign trade was liberalized and the economy completely opened to the outside world for foreign trade and investment.

As the reforms took effects, the economy started to grow in the beginning of 1990s. High rates of economic growth over have been recorded the last 15 years. The country has become the second biggest rice exporter in the world. Vietnam has been considered as a potential and prospect destination for foreign investment with macroeconomic and political stability and other advantageous conditions. As a consequence, people's incomes almost tripled over the last ten years and the living standard of the people substantially lifted.

Achievements recorded during over the last 20 years of economic reforms in Vietnam had tied down to all aspects of the socio-economic changes, and therefore there is no way for Vietnam to drive away from a market oriented economy mechanism. Hopefully, new significant changes in policy direction and reform implementation toward more openness would be made in Vietnam in years to come in order to switch the economy toward complete market system for freer mechanism and higher growth. The prevailing institutional limitations and constraints mean that there is still a large space for Vietnam to change in order to meet expectation of the initial reforms as to create a stabile and strong market economy in its socialist orientation.

Chapter three
Multidimensional methods to assess welfare inequality

3.1. Introduction

Economic welfare and its distribution across the population of an economy are of crucial importance to the full understanding of the impacts of induced economic change (Maasoumi and Nickelburg 1988). The analysis of welfare inequality has taken on a central role in the economics literature (Atkinson 1996). Conventionally, the methodology approached by most of the traditional studies in analysis of welfare inequality is to apply a single-dimensional welfare indicator such as income/expenditure or other forms of material wealth like housing and durables. However, single dimensional welfare as income/expenditure may not sufficiently and adequately reflect the level of social and economic welfare in a given society/country, and it even may display unsatisfactory properties of welfare inequality. The single dimensional approach may also not be able to capture all aspects of welfare inequality, whereas welfare state depends not only on income but also on other welfare attributes like healthcare, education, and housing. In addition, education and healthcare themselves are multidimensional welfare components. They are constituted by more than one indicator. Thus, using a single dimensional approach to inequality analysis for these welfare components may not be appropriate and would bring about misinterpretation of welfare inequality and would also provide short of basic information on welfare inequality. In his various works on welfare and inequality, Sen (1981, 1985, and 1992) pointed out that poverty and welfare inequalities are multidimensional issues. Maasoumi (1989) also stated in his study of inequality using income, basic needs and physical quality of life, common inequality measured by income of an individual may have exaggerated the level of welfare inequality as his results showed that multi-attribute inequality is lower than income inequality. Therefore, distribution patterns of income using a single indicator may not fully reflect the individual's welfare situation as well as a his needs or abilities, particularly those that could not be priced as they are not valued in monetary term(s) such as education, healthcare and basic needs (Sen, 1985, 1997). Instead, the measurement of inequality in income or expenditures would aggravate or moderate the level of welfare inequality in a society, and would provide a wrong indication about welfare inequality to people concerned and policy makers.

Recently, there has been an increase in studies on multidimensional welfare distribution. The United Nation Human Development Index (HDI), which combines indicators of purchasing power parity (PPP) per capita GDP, life expectancy at birth, and literacy rate to calculate a composite index presenting people's living standards, is perhaps the best-known multidimensional index and most widely applied in present researches on welfare inequality. The multidimensional approach to welfare inequality becomes more popular in empirical studies since the multidimensional inequality approach provides more adequate information on people's welfare inequality and would be able to capture the major properties of people's welfare distributions. In addition, the multidimensional approach is considered as superior to the traditional (unidimensional) approach because it is recognized that the aggregate attributes are considered as "ideal" indices for evaluations of individual welfare when the distribution of welfare is the primary concern of analyses (Maasoumi 1986).

Over the past 20 years, the economic reform (*doi moi)* in Vietnam has brought about the improvement in living standards of its people. However, household expenditure and welfare distribution in the country have been moderately deteriorated (World Bank 2004). Many studies jointly conducted by the GSO of Vietnam and the World Bank's Vietnam Office, show that household expenditure distribution in Vietnam has been deteriorated over time as Gini index increased from 0.34 in 1993 to 0.35 in 1998, to 0.37 in 2002, and to 0.41 in 2006. However, there is still a lack of systematic studies that look into the matter using the multidimensional approach as the people's welfare is of a multidimensional nature. In fact, most researches on welfare inequality so far conducted in Vietnam employed a conventional approach (using either Theil index or Gini coefficient) to measure and analyze the household expenditure distribution, while the distributions of other welfare attributes such as education, health care and other basic needs have not received full attention. Hence, inequality in the country would be seen in relation to the lack of important welfare attributes like education and health. Therefore, general facts about welfare distribution in the country are still short since the expenditure distribution may not provide adequate information on welfare distribution. Moreover, expenditure inequality may either aggravate or mitigate the inequality in the distribution of welfare in the country. To overcome this objection, my study aims to analyze multidimensional regional welfare inequality in Vietnam.

The nature of the multidimensional method of inequality measurement and analysis lies in the way multidimensional aspects of welfare should be measured and aggregated into a composite index by aggregating a number of welfare indicators. For this purpose, this study employs the one-stage Theil decomposition method. It considers the three-level hierarchical structure of a country: region -> province -> district as displayed in Figure 3.1, and decomposes the total regional inequality, as measured by a Theil index based on a provincial level composite index of welfare components, into two components: the between-province and within-province inequality.

The structure of this chapter is as follows. Following this introduction, a brief review of the literature on regional welfare inequality is provided in section 3.2. In section 3.3 we present the method for measuring and analyzing the multidimensional regional welfare inequality. In order to use the method the data for analysis must be compatible, so the data should be transformed. The discussion of data transformation is done in section 3.4. The overview of the procedure for weights and composite index estimation, and then the calculation of Theil's index is structured in section 3.5. Finally, the chapter's brief summary is provided in section 3.6.

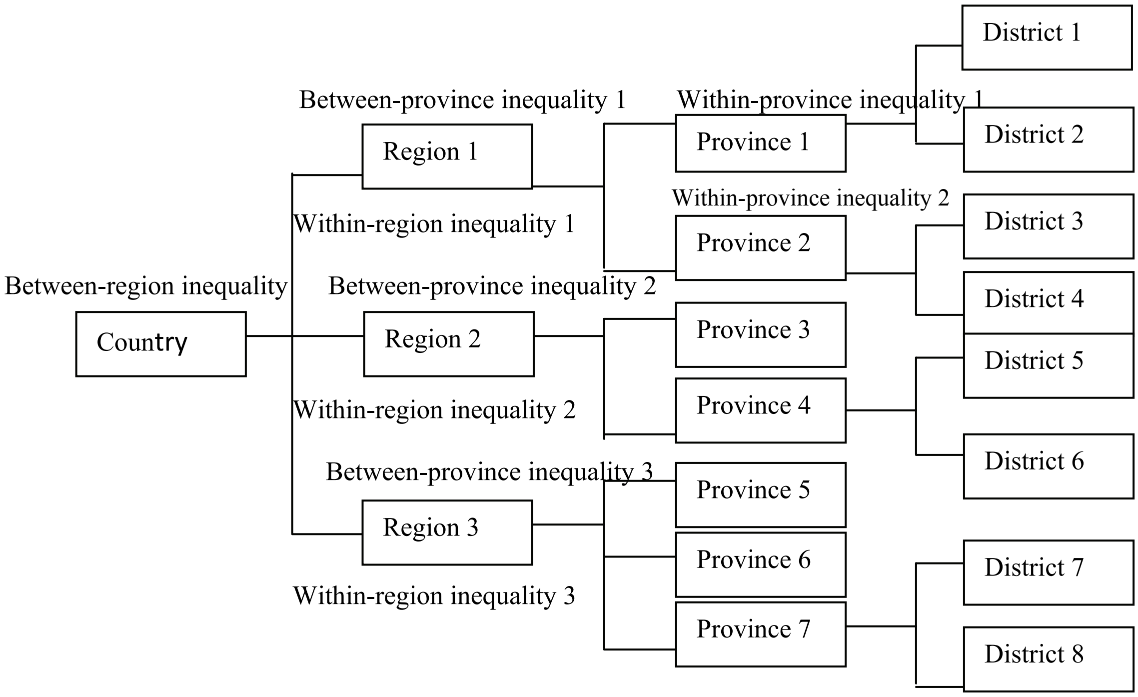


Figure 3.1. Three level hierarchical structure: region-province-district

3.2. A brief review of literature on multidimensional regional welfare inequality

One of the most important issues in development economics is the welfare inequality and its relation to poverty and development. Therefore, the issue has induced many researchers to conduct research on the issue that resulted in a large volume of literature. In the very beginning, most researchers in the field paid attention to, as aforementioned, income or expenditures inequality. Recently, multidimensional aspects have been brought to the fore since a single dimensional approach had been found to have limited significance in reflecting people's welfare distribution (Sen, 1991, Maasoumi, 1986).

Theoretically, one of the most pioneering researches on welfare inequality using a multidimensional approach is the work by Fisher (1956) who developed the idea of a multidimensional distribution matrix. After Fisher, more recently there was a study by Kolm (1977) who conducted a theoretical study on multidimensional egalitarianism. In this paper he is the first to discuss and to lay the concept of multidimensional inequality as he pointed out that multidimensional inequality should be a function of a uniform inequality of a multivariate distribution of goods or attributes. Meanwhile, according to Atkinson and Bourguignon (1982), multidimensional inequality should be a function of the cross-correlation between distributions of goods or attributes in different dimensions.

Studies in which inequality in more than one welfare variable is analyzed are still sparse compared to the ones based on the single approach. In this chapter, we try to link to the most relevant empirical studies in the field.

Using the aggregate function developed by Ram (1982) to compute the composite index in their research on multidimensional inequality for 120 countries for three welfare components namely Physical Quantity of Life Indicator (PQLI)[12], Basic Needs Indicators (BNI)[13], and per capita GDP, Maasoumi and Jeong (1985) found that the multidimensional welfare inequality is more serious in centrally planned countries than in market economies while income inequality in these countries was known to be more equal. Also, in this study the authors found that the size of multidimensional welfare inequality is less severe than the one of income inequality within a country.

In their studies, Lydall and Lansing (1959), Projector and Weiss (1966) and Toussig (1973, 1976) have found that indicators of economic welfares are distributed differently. This finding has drawn support to the idea of diversity in the distribution of economic welfare indicators. The diversification is considered as basic background for analyzing multidimensional welfare inequality.

Other studies on multidimensional welfare inequality conducted so far found the reverse trend of multidimensional inequality against income/expenditure inequality. In most cases, multidimensional inequality seems to be less serious than univariate inequality whereas only one indicator is considered in measuring the welfare distribution (amongst others, see Maasoumi, 1989, Maasoumi and Nickelsburg, 1988; Quadrado, 1999; Lugo, 2004; and Justino *et al.,* 2004).

In Vietnam, many research projects have been carried out to study the change and the pattern of welfare distribution over time. However, most of these studies had mainly focused on household expenditures inequality. Among many others, these works done by Dollar *et a*l.(1998); Binh Nguyen *et al.* (2003); Minot *et al.* (2003) and a number of other researches by Heltberg (2003); Liu (2001), GS); and the World Bank (1995, 2000, 2002 and 2004). These studies used traditional inequality index (Gini) to measure and analyze overall inequality and regional inequality by using the data from the Household Living Standard Survey in 1992 or 1997. These studies found that accompanied with the high rate of economic growth during nearly two decades, the household expenditure inequality has been recorded to moderately increase over time.

The fact is that, despite the number of studies on welfare distribution conducted in Vietnam so far, researches on multidimensional welfare inequality are absent. Dollar *et al.* (1998), in their research entitled "Household Welfare and Vietnam's Transition" had mentioned about inequality in education and healthcare in the country. However, in their study they just take each welfare attribute as a single indicator to investigate the differences of the distribution of these welfare components among regions in Vietnam. What the research lacks is a comprehensive approach using the multi-indicator method or aggregate function to investigate the distribution status of each welfare component separately. Therefore, the findings may still underestimate or overestimate the regional inequality in Vietnam.

[12] The PQLI indicators consist of life expectancy index, infant mortality index, and adult literacy rate.

[13] The BNI includes adult literacy, life expectancy, safe water access, physician supply index and calorie intake index.

3.3. A method for measuring and analyzing multidimensional regional welfare inequality

In section 3.3, we discuss the method for measurement and analysis of multidimensional welfare inequality based on the work by Maasoumi (1986), and the estimation method developed by Flury (1988). The structure of this section is as follows. The theoretical approach to the measurement of multidimensional inequality will be provided in section 3.3.1. In section 3.3.2 we elaborate the estimation models based on the common principal component model (PCA), and partial common principal component model (PCPC) developed by Flury (1988).

3.3.1. Theoretical approach to measurement of multidimensional inequality

As discussed in section 3.2**,** different indicators of economic welfare are distributed very differently. Without this diversity in distribution there would be no need for a multidimensional measure (Maasoumi, 1988). Indeed, the measure of multidimensional welfare depends on this diversity. Therefore, measurements of multidimensional inequality are applied when multiple indicators[14] of a welfare component are included in the measuring process. These multiple indicators can be transformed into a "one-dimension indicator" (or a single indicator) in the form of composite index using aggregate function by Maasoumi (1986), see Appendix 3.1. This aggregate function is as a utility-like function of all the attributes received (Maasoumi 1986). In order to measure multidimensional inequality, we follow the work done by Maasoumi (1986) in which he introduced the use of the Theil index to measure regional welfare inequality, and the method of the ordinary one stage Theil decomposition used by Anand (1983)[15]. The latter is extended for use by Akita (2001).

Consider the following hierarchical structure of a country: region – province. Using province as the underlying regional unit, according to Maasoumi (1986) and Anand (1983), total regional welfare inequality can be measured by the following Theil's second index.

$$T_{TT} = \sum_i \sum_j \left(\frac{W_{ij}}{W} \right) \log \left(\frac{W_{ij} / W}{N_{ij} / N} \right) \qquad (3.1)$$

where:

W_{ij} is the composite index of the welfare under consideration of province j in region i,

W is total composite index of the welfare of all provinces, and is defined as follows:

$$W = \sum_i \sum_j W_{ij} \qquad (3.2)$$

[14] Multiple indicators hereby is understood as the number of indicators to be considered larger than one, i.e. there must be at least two indicators be included into measurement.

[15] This is the one-stage decomposition of Theil indexes, which will be applied in our study. In his study, Anand (1983) computed Theil's index for income. We extend the method to compute Theil's second measure to the case of other welfare components such as health and education which are constituted of multiple indicators.

N_{ij} is the population of province *j* in region *i*, and
N is the total population of all provinces, and is defined as follows:

$$N = \sum_i \sum_j N_{ij} \tag{3.3}$$

If we denote T_{bpi} as between-province inequality for region *i*, and define as follows:

$$T_{bpi} = \sum_j \left(\frac{W_{ij}}{W_i} \right) \log \left(\frac{W_{ij}/Wi}{N_{ij}/N_i} \right) \tag{3.4}$$

then total Theil's second index in equation (3.1) can be decomposed into

$$\begin{aligned} &= \sum_i \left(\frac{W_i}{W} \right) T_{bpi} + \sum_i \left(\frac{W_i}{W} \right) \log \left(\frac{W_i/W}{N_i/N} \right) \\ T_{TT} &= \sum_i \left(\frac{W_i}{W} \right) T_{bpi} + T_{BR} \\ &= T_{WR} + T_{BR} \end{aligned} \tag{3.5}$$

where:

W_i is composite index of the welfare computed for region *i*, and defined as:

$$W_i = \sum_j W_{ij} \tag{3.6}$$

N_i is the total population of region *i*, and defined as:

$$N_i = \sum_j N_{ij} \tag{3.7}$$, and

T_{BR} denotes the between-region inequality.

$$T_{BR} = \sum_i \left(\frac{W_i}{W} \right) \log \left(\frac{W_i/W}{N_i/N} \right) \tag{3.8}$$

Equation (3.8) is the ordinary one-stage Theil's second inequality decomposition, in which the total inequality T_{TT} is the sum of the within-region component (T_{WR}) and the between-region component (T_{BR}). As we observe from equation (3.8) the within-region component is a weighted average of the between-province inequalities for each region (T_{bpi}).

As we observe from the equations in the previous section, in order to arrive at the Theil's second index of inequality, there is a need to calculate the composite index of the welfare component under consideration. Both equations (3.2) and (3.6) above are used to determine the composite index of a welfare component under consideration. Equation (3.2) presents total value of composite index of the considered welfare for the whole country while equation (3.6) determines the total composite index of the considered welfare for the region *i.* As we observe from both equations, both the composite index at the country level (*W*) and the composite index at the regional level (W_i) depend on the composite index at provincial level (W_{ij}). Because our aim is to consider a multidimensional welfare which implies that a welfare component under consideration is constituted by multi indicators (s_{ij}). Thus, the composite index at the provincial level of the considered welfare component presented in equations (3.2) and (3.6) (W_{ij}) is defined as follows (Maasoumi, 1986 and Quadrado, 2001):

$$W_{ij} = W(s_{ijm}) \tag{3.9}$$

where s_{ijm} is a vector of indicators of the welfare under consideration of province *j* in region *i* as, for example, number of primary teachers per 1000 primary students, number of primary schools per 1000 primary students, number of lower secondary teachers per 1000 secondary students, number of lower secondary schools per 1000 secondary students, number of upper secondary teachers per 1000 secondary students and number of upper secondary schools per 1000 secondary students etc. in province *j* of region *i* in the analysis of education facility component or number of hospital beds per 1000 population and number of doctors per thousand population, proportion of immunized children, number of registered drug chemists per 100,000 population etc. in province *j* of region *i* in the analysis of healthcare facility component. Since the multidimensional nature of welfare components is considered in this study, the welfare components consist of s_{ijm} indicators, where subscript *m* denotes the number of indicators (variables) considered in the corresponding welfare component (*m=1,..,p*), for example, as in the case of the measurement and analysis of multidimensional regional inequality for the education facility component which includes 6 indicators (or variables) (see above), so we have $m = 6$.

To compute the composite index in equation (3.9), we use the aggregate function. According to Maasoumi (1986) the aggregate function for composite index of a welfare component for the j^{th} province is defined as follows[16]:

$$w_{ij} = \sum_{m=1}^{p} \delta_{jm} s_{ijm} \tag{3.10}$$

where:

δ_{jm} are the weights associated with the s_{ijm} indicators,

s_{ijm} are the selected indicators for the welfare component under analysis.

The estimation of the composite index defined in equation (3.10) requires the weights (δ_{jm}) to be estimated. We will discuss the method of estimating these weights and computing composite index later in the study in section (4.3.2) that presents the measurement models.

The reason behind the selection of the Theil's second index is that the Theil's second index satisfies the property of additive decomposability and it is being very informative about the status of the welfare distribution (Maasoumi 1986). It also satisfies several desirable properties as a measure of regional income (welfare) inequality, i.e., mean independence, population-size independence, and the Pigou-Dalton principle of transfers (Bourguignon 1979; Shorrocks 1980). Its property of additive decomposability means that overall inequality can be written as the sum of within-group and between-group inequality. This property is very useful for understanding magnitudes and trends of each component of total inequality, and analyses of components attributed to total inequality. Mean dependence implies that the inequality index remains unchanged if every region's income is

[16] For more details of this aggregate function, please refer to Appendix 3.1 at the end of this chapter.

changed by the same proportion. Population-size independence implies that the index is unchanged if the number of people in each region is changed by the same proportion. This means that the index depends only on the relative population frequencies at each region, not the absolute population frequencies. Lastly, the Pigou-Dalton principle of transfer means that any income (or other welfare in a broader sense) transfers from a richer to a poorer region that does not reverse their relative rank in that income (or welfare) would reduce the value of the inequality index. In addition, Theil's second measure is also able to provide an overall view of the welfare distribution.

3.3.2. Estimation models

To compute the composite index (W_{ij}) presented in equation (3.10) we need to have the coefficients of indicators (s_{ijm}) and their associated weights (δ_{jm}). The coefficients of indicators are computed from the data, but for their associated weights, we need an estimate using appropriate models. To do so, we use either the Partial Common Principal Component Model (PCPC) or the Principal Component Models (PCA) wherever appropriate, see section 3.3.4. The details of the PCA model and PCPC model are given in Appendix 3.2 and Appendix 3.3 respectively at the end of this chapter. The PCA is used in case the PCPC model does not fit the data[17] (the procedure of estimation and test using either PCPC model or PCA model is discussed more extensively in Section 3.5). The weights drawn from the estimates using either PCPC or PCA are used to compute the composite index at the provincial level presented in equation (3.10) . Then we use these weights to compute weights for the higher levels (at the regional and national level) presented in equations (3.6) and equation (3.2) respectively. The PCPC model is known as a generalisation of the classical Principal Component Analysis (PCA). For more details about the method, we refer to Flury (1988). In this section, we will briefly discuss the estimation method using the PCA model and the PCPC model.

The section is structured as follows. The discussion of the estimation method using the PCPC model is given in section (3.3.2.1). In section (3.3.2.2) we discuss the estimation method using the PCA model.

3.3.2.1. Estimation using the PCPC model

As we have already mentioned above, the purpose of the PCPC method is to obtain common weights associated with the coefficients of indicators of the welfare component under consideration for composite index computation. The PCPC model developed by Flury (1988) is a generalization of the Principal Component Analysis (PCA), which has been an appropriate device for longitudinal analysis. It has been applied by many researchers to calculate composite index in inequality analysis using Theil measures (see, among others, Zandvakili (1999, 1992), Maasoumi and Jeong (1985)). The nature of the PCPC method[18] is to compare the structure of two or more covariance matrices in a hierarchy of similarities. The hierarchy of similarities is built upon the realization that covariance matrices can share more complex relationships between one another than just being equal or not. For example, one matrix might be identical to another except that each

[17], We use the Maximum Likelihood Test to test for the fit of the PCPC model
[18] For more detailed about the use of the PCPC model and estimation software, see Phillips (1997).

element of the matrix is multiplied by a single constant. We would then say that the matrices are proportional to one another. A more precise definition of proportionality is that the matrices share identical eigenvectors (or principal components), but their eigenvalues differ by a proportional constant. This suggests that another relationship between matrices could be that they share principal components in common, but their eigenvalues are different (the Common Principal Components or CPC model). Similarity, the matrices could share one, two, or up to *p-2* (where we are working with *p x p* sized matrices) principal components in common out of the *p* total possible components (*p* is number of indicators or variables of the welfare component under consideration). This is the Partial Common Principal Components model. The PCPC model stops at *p-2* components because, since the principal components are defined to be orthogonal to one another, if you know *p-1* of the components, you necessarily know the final one, leaving you with the full CPC model.

The formal PCPC model is a data reduction technique of a large data set with *t* samples. Thus, the number of variables *p* measured over several samples needs to be reduced to a subset of *q* components ($q<p$) that is associated with the largest eigenvalues. The basic assumption in the PCPC model is that the first component of all *t* samples ($t=1,\ldots k$) is identical whereas the remaining components are specific to each sample (t=3 in our study). Because the PCPC method is basically based on the PCA method, thus the technical estimation and interpretation of the principal component (in PCA) or common principal component (in PCPC) are the same. For more detailed information about the PCA method, we refer to section 3.3.2.2 below. Because the main interest of our study is to calculate a composite index of all the variables under consideration in one welfare component in a single index for Theil's second measure computation, only the first common principal component is used. Thus, *q* in this study equals one.

The reasons behind choosing the PCPC technique are as follows. As described by Flury (1988) the method of PCPC analysis is appropriate for the longitudinal analysis. Meanwhile, the data used in this study are longitudinal. In addition, the multivariate models as PCPC and PCA can be used for longitudinal analysis to calculate composite index as we have seen the PCA method is widely applied in many studies by the World Bank to calculate Human Development Index and multidimensional poverty index. Thus, we found that these models are appropriate to be applied here.

3.3.2.2. Estimation using the PCA model

Principal Component Analysis (PCA) is used when the PCPC model does not fit the data in our analysis (the Maximum Likelihood test for common principal component for all the samples is rejected). In other words, the PCA model is applied when the component coefficients of the *t* samples are specific to each sample (details are provided in section 3.5 below). In this case, PCA is considered as an appropriate method for estimating coefficients of indicators or weights associated with indicators of the welfare under consideration.

PCA is known as a technique that is used to reduce multidimensional data sets to fewer dimensions for analysis by retaining those characteristics of the data set that contribute most to its variance, by keeping the principal lower-order components and ignoring higher-order ones. Such lower-order components often contain the most important aspect of the data. The multivariate method of computing principal components is a relatively straightforward device for parsimonious representation of a multiplicity of related variables for any given observation set[19]. The method essentially consists of computing such linear combinations of the original variables that capture successively the largest proportion of the variance in the original variables. Thus the first principal component is that linear combination which captures or explains the highest fraction of the variance in the original variables; the second explains the largest part of the remaining variance, and so on. The various principal series are orthogonal to each other. Technically speaking, the principal series are constructed by taking the sum of the original variables, the weights being the elements of the characteristics vectors of the covariance matrix of the original variables, and the characteristics being arranged in the descending order of the size of their associated characteristic roots. The weights in each principal component are unique up to a scale factor and are determined from the data so as they could capture the maximum variance of the original variables.

Finally, PCA is also used as a tool in exploratory data analysis and for making predictive models. For example, if we use the PCA method to analyze a sample of n observations that contains p correlated variables (or indicators), we would have p principal components of the data for analysis. Usually, first q component ($q<p$) with lower-order that are linear combinations of the original variables having maximum variance, are those containing most information from the data, so we rely on these components only for analysis. The principal components derived for this study are based on the covariance matrix of the relevant variables, and are literally weighted average of the underlying variables. Because the main interest of this study is to calculate a composite index of all the variables under consideration in one welfare component in a single index for Theil's second measure computation, only the first principal component is given. Thus, q in this study equals one.

The discussion above indicates that PCA is a rather natural, straightforward, and parsimonious mode of representation of observation sets with many related variables. Whatever may be the number of original variables, they can be well represented by a much smaller number, possibly by just one variable in the form of a principal component. These components can often capture a very large faction of the variance in the original data set and can thus serve virtually the same function as the full set of original variables, but in a much more compact form.

3.4. Method of data transformation

One of the purposes of our study is to rank provinces in term of composite indices. The ranking orders with respect to composite index would help us figuring out favored provinces with respect to a welfare component under consideration. It means, in this context, the provinces having a higher composite index are considered more favored than those having a smaller composite index. Therefore, all selected indicators are expected to be positive. If they are positive they would have positive impact on the multidimensional regional distribution of a welfare component. Thus, for an indicator like infant mortality rate in the health status analysis that has a negative impact on the composite index (hereafter we call these indicators the negative indicators) while another indicator like number of doctors per 1000 population has a positive impact on the composite index

[19] The method is described in most standard texts on multivariate analysis. The discussion here is largely based on chapter two and three of Flury (1988) and chapter seven of Morrison (1967).

(hereafter we call them the positive indicators), we should transform them all to have a positive impact on the composite index.

In addition, these indicators should be uniform and commensurable in term of measuring unit in order to make use of the principal component method (either PCPC or PCA models) for deriving composite indices for welfare components under consideration. To do so, we need to transform and rescale all indicators to make their respective values lie in the range between 0 and 1. It is not problematic for variables measured in term of percentage like infant mortality and childhood immunization because their values were already in the range between 0 and 1. However, for the variables, those measured in the terms of quantity such as hospital beds, medical doctors and pharmacists we need to rescale them in order to make their values lie into the range between 0 and 1. It is not easy to rescale the variables to a 0-1 scale except by making a transformation of the type done by Morris (1979). Such a transformation can, however, greatly alter the relative ratios across provinces. Therefore such a scaling does not seem quite desirable, and there appears some merit in preserving the relative ratios as much as one can. To do so, we need to equate the highest value to 1, the values of indexes equal 0 and 1 for smallest value and highest value respectively. Then we use the following formulas developed by Nijkamp (1978, 1988) with reference to Morris (1979) to transform the data to be in the 0-1 scale.

$$s^{i}_{ijm} = \frac{s_{ijm}}{\max(s_{ij})} \text{ if } s_{ijm} \text{ is a positive indicator} \qquad (3.11)$$

$$s^{i}_{ijm} = 1 - \frac{s_{ijm}}{\max(s_{ij})} \text{ if } s_{ijm} \text{ is a negative indicator} \qquad (3.12)$$

where:

s^{i}_{ijm} is the transformed index value of the *m-th* indicator of welfare component under consideration of province *j* of region *i*,

s_{ijm} is the value of the *m-th* indicator of welfare component under consideration of province *j* of region *i*,

$\max(s_{ij})$ is the highest value of the *m-th* indicator of welfare component under consideration of province *j* of region *i*.

The results of the transformation are the data in the form of index values of variables. These index values were used as input for either the PCPC or PCA model estimation.

3.5. An overview of the estimation procedure

As we have discussed earlier in this chapter, to compute Theil's second measure of the multidimensional regional inequality we need to determine weights attached to indicators. These weights would be common weights or distinct weights depending on whether the data fit the PCPC model or PCA model respectively. The following steps are to be followed to determine these weights, and then the composite index and Theil's second measures or the empirical analyses.

Step 1: Covariance matrix of indicators for each period is derived using STATA Software. This matrix is used as input for the PCPC model in the next steps,

Step 2: PCPC method is applied for all the three periods (*t=3*) under consideration to estimate common component coefficients of indicators using inputs as covariance matrices of variables obtained from Step 1. The output includes maximum likelihood estimates for each period, a test for common component, eigenvalues and eigenvectors, etc..

Step 3: In order to examine whether the PCPC model fits the data, we resort to the Maximum Likelihood Test to test the null hypothesis $H_{PCPC}(q)$ (q=1: first common component) of a common component for all the three periods (refer to Step 2) against the alternative H_{PCA} of a distinct component for each period (refer to Step 8). If the test is accepted we go on to step 4 and then step 5. If not, we go to step 6.

Step 4: Together with the test conducted in Step 3, we check the proportion of variance captured from the original variables by the first common component obtained from estimating the PCPC model (Step 2) (whether it is as high as at least 60% as pointed out by Flury (1988)), the standard errors of the component coefficients of the indicators (whether it is smaller than 10% of the value of the variable coefficient) and the covariance and correlation matrix of the Partial CPCs for all the three periods (no substantial correlation between partial common principal components),

Step 5: If the null hypothesis in Step 3 is not rejected and all the checked criteria mentioned in Step 4 are met, then the indicators' coefficients estimated in step 2 will be interpreted as robust and reasonable for interpretation. These coefficients are just a reflection of the high proportion of the variance in the original variables recovered by the first common component, and will be used for interpretation and determination of the weights. Then, we go on to step 9.

Step 6: If the null hypothesis in Step 3 is rejected for all the three samples, then it is possible to reduce the number of periods t to $t\text{-}r$ ($2 \le r \le t$) until the test for the null hypothesis of a common component is not rejected for the number of t samples (in our case $t=3$ and $r=1$). Then for the $t\text{-}r$ samples with the first common principal component we check the proportion of variance recovered together with standard errors of the component's coefficients and the covariance and correlation matrices of the partial CPC for these two periods (refer to step 2). If these checks are satisfactory, then the first common component's coefficients will be interpreted as robust and significant for these t - r (=2 in our study) periods (refer to Step 4). For the remaining r periods (r = 1 in our study) that the PCPC did not fit the data, we will use the PCA method to determine coefficients of indicators (see Step 5),

Step 7: If all the tests conducted in Step 3 and Step 6 are rejected. This means that it is not possible to reduce the number of parameters estimated whereas the PCPC model does not fit the data. If there is no any common component shared by any two or more periods (samples), we resort to the PCA method (go to step 8) to determine the indicators' coefficients, which are exactly the first principal component's coefficients. In this case, the weights attached to indicators are distinct and not the same for all the three periods.

Step 8: The PCA method is applied to estimate component coefficients for each period separately. The estimates include the components' coefficients of indicators, variance and correlation matrix for each period and proportion of variance by each retaining components. We use component coefficients estimated by the PCA to determine weights, then go on to step 9,

Step 9: When all those weights are determined, composite indices are computed using equation (3.10), then equation (3.6), equation (3.2). Finally, Theil's second measures are computed using equations (3.4), (3.5) and (3.8), and the spatial inequality is analyzed under these Theil's second measures.

3.6. Summary

In chapter three we provided a short introduction to the multivariate approach to welfare distributions, a brief review of related literature, the model construction and the data reduction to be applied for the analysis. In the model construction, we use Theil's second measures for multidimensional welfare inequality, and its decomposition into between and within terms of inequality in order to distinguish the between and within regional inequality in the empirical analysis. In order to compute Theil's second measures of inequality, there is a need of estimating a composite index based on the aggregate function by Maasoumi, in which the indicators' coefficients and their associated weights need to be estimated. To do so, we use the principal component models (PCA) and partial common principal component models (PCPC) developed by Flury. These two types of models are discussed extensively in order to help the reader to comprehend the subsequent empirical parts of the thesis.

Chapter four
Healthcare Facilities and Health Status

4.1. Introduction

Like many other important dimensions of welfare as income, education, housing, food and nutrition, health plays an important role in the development process. Health is believed to have both a direct and indirect effects on people's well-being and the development of a country as well. Health is both consumption and production. Health also affects social impacts and productivity of the people. Like well-being, health itself is also a multidimensional concept. Its components are health status of the people and healthcare services, which are mutually correlated. Each component of health comprises multiple indicators. This is the reason why health should be considered using a multi-dimensional approach, and the correlation among components of health should be examined and analyzed.

Healthcare services contribute to improving health status of people. In its turn, health status may contribute largely to development. Health status may shape the opportunities people face for future progress and achievement because people with good health would have a more joyful life, more opportunities and they are more productive than those with worse health (World Bank, 2006). In addition, a good health status may not only mitigate the budgetary and financial burden on society but also create more benefits to such a society as good health could increase people's productivity and innovation. As in the capability approach to studying well-being, Sen argued that in welfare economics exclusively focusing on income inequality would be a fatal mistake, because there are more aspects rather than income that people have for leading a valuable life, and on their capabilities to undertake activities in their daily life (Sen 1985, 1987a, 1987b, 1992, 1993, 1998). Those activities would be reading, working or being politically active, or enjoying positive states of being, such as being healthy or literate could help people achieve more in their lives.

Recent evidence shows that individuals/groups were facing different opportunities to get access to health facilities and services in order to attain a good health status (World Bank, 2006). These unequal opportunities may be brought by unequal distribution of healthcare facilities, unequal allocation of resources for healthcare and other healthcare programs such as preventive and cure programs as well as differences in geographical factors among regions.

Recognizing its importance to development, making a better healthcare system and then striving to achieve a better health status for people were major goals that most governments in the world, especially in the developing world aimed to achieve. It is also a big challenge during the development processes. Providing better healthcare programs and facilities by improving health facilities and services therefore are priorities in the development program of the most governments because providing good and adequate healthcare facilities would make living better (World Bank, 2006). However, the problem that challenges most developing countries is that there is an existence of inequality in the distribution of health status and allocation of healthcare facilities among groups and regions. This tendency is problematic because it may broaden the gap among regions within a country. Thus, a study on spatial inequality in healthcare facilities and health status, therefore, becomes very important. The findings of the study are important to policy makers for appropriate healthcare policies towards sustainable development because it would provide justification and a reference source for the policy making process.

Since the implementation of the reforms (*doi moi)* in 1986, Vietnam has attained good achievements in sustaining a high rate of economic growth for over the last 15 years (5-9% per annum). Meanwhile, the country has also recorded an impressive decline in poverty (measured in head count index) over the time, from 73% in 1991 to 19% in 2004. The *doi moi* has also brought about improvement in people's welfare in general and people's health status and healthcare facilities in particular. However, there were evidences about disparities in health status and healthcare facilities among regions in the country (see Table 4.1). Health status has not been the same, and health facilities have not equally distributed among people and among regions in the country as the budget spent on health was regionally unequal. Differences in lifestyles and habits as well as people's awareness among regions and ethnic groups also cause different health status among regions and people groups.

The organization of this chapter is as follows. Following this introduction, a brief overview of the healthcare structure is provided in section 4.2. Section 4.3 is scheduled for an elaboration of the health status and healthcare policies before and after the *doi moi*. This section is followed by a brief description of data and the elaboration of the choice of indicators in section 4.4. Spatial interrelation and inclusion of contiguity in health facilities will be discussed in section 4.5. In section 4.6, the results of the analyses will be reported and discussed. The test for spatial dependence among health dimensions and analysis of spatial interrelation will be reported in section 4.7. Finally, a brief summary of the chapter is given in section 4.8.

4.2. An overview of current structure of health services in Vietnam

The North Vietnam adopted a planned and centralized system after the country's victory over France in the Dien Bien Phu Battlefield, which led to the Geneva agreements in 1954. The country's healthcare system was hierarchically structured and bureaucratically organized (Figure 4.1). Under the system, the Ministry of Health (MOH) is a governing body. The implementation of healthcare activities were carried out in the local governments at the provincial, district levels, at the village and ward levels controlled and monitored by the MOH. In addition, the finance for healthcare was subsidized under the centralized system. There was not a system of health insurance in the past, but people have been provided the services and medication for free since the independence in 1945 up to the inception of the *doi moi* in 1986.

The overall structure of the healthcare system in Vietnam is as follows. The central governing body (MOH) of the health system governs and manages the system vertically over 61 provinces and urban centers including five municipalities: Hai Phong, Hanoi, Da Nang, Ho Chi Minh City and Can Tho (Appendix 4.1)[20]. The highest level of the health system, the MOH governs almost all the health services and healthcare development, including controlling eight medical schools and overseeing the education of health workers,

[20] Appendix 4.1 contains 61 provinces and urban centres in Vietnam. However, the administrative division has been currently changed as the number of provinces and urban centres increased to 64 with the inclusion of Dien Bien (separated from Lai Chau), Dac Nong (separated from Dak Lak) and Binh Phuoc (separated from Song Be). Because the data is only available for the old administrative system, this thesis uses the old administrative division for the analysis.

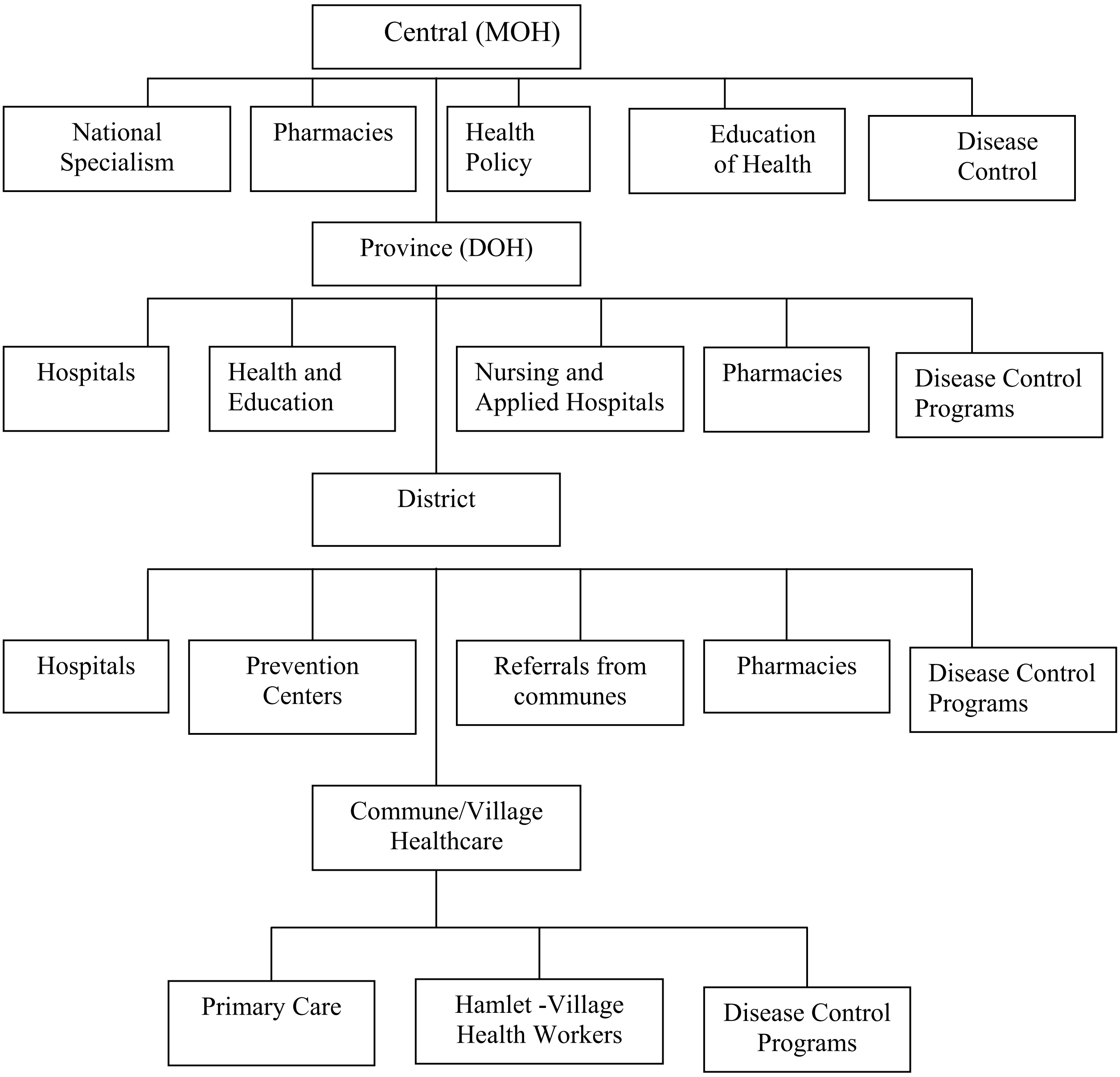
pharmaceuticals (including manufacturing, importing, regulating and distributing), taking in charge of health policy making process, and coordination of the National Health Budget. Being the two largest urban centers in the country, Hanoi and Ho Chi Minh City host many first-class facilities and national specialism institutes. The MOH also governs nearly 11,000 health stations, around 1,300 inter-communal clinics, some 720 general hospitals, 80 specialized hospitals, and many others specialized clinics and health centers. With the process of reforms, private and foreign clinics have been established in Vietnam recently. As a result, health services have been improved as the person-doctor ratio increased from 2380:1 in 1995 to 1960:1 in 2004 (Vietnam Academy of Social Sciences, 2006).

At the lower level, the provincial hospitals and clinics can be found at the provincial capitals. These healthcare bases provide all the major healthcare services at the provincial level such as health consultation, in-patient treatment, implementation of various disease preventive programs, and immunization schemes for children and the like. The number of districts in each province varies from 8 to 20 on average with differences in population. There is a provincial health service department called Department of Health (DOH) in each province which plays a role as the governing body for the provincial health system. The DOH is financed by the central government through MOH for its function and operation. The DOH is responsible for manufacturing, importing and distributing drugs, and coordination of treatment and preventive activities in areas such as malaria, tuberculosis, leprosy, AIDS, immunization for children and other diseases at provincial level. The provincial hospitals provide laboratory services. It is also a training centre for nurses, midwives and assistant doctors.

At the district level there are district hospitals with fewer specialism compared to the hospital at provincial and central levels. In addition, the size of the district hospital is usually smaller (the average size is of around 100 to 200 beds). Prevention centers and pharmaceutical distribution services for the district level or lower (for example, village or hamlet and ward (a division of a city or town that is lower than district level) can also be found at this level.

At the village level, there may be one doctor working in a village clinic. Usually, health workers at the village level are assistant doctors, midwives, pharmacists, traditionally medical practitioners and nurses. Most common health services in the village level are an outpatient care for simple diseases, minor surgery, assistance to normal deliveries, and supervision to health workers working at the hamlet level. Usually, there is one trained health worker in each hamlet. The nurse and health workers provide healthcare services including disease diagnosis and treatment, coordinate home care, and provide caring services to infectious diseases, mental health and postpartum patients in villages and hamlets.

There are also Red Cross practitioners working in the village and hamlet guided by a nurse. The Vietnamese Red Cross is different from the international Red Cross as it is a voluntary organization following the hierarchical structure as of the healthcare system. It is an independent organization working closely with the MOH at all levels to partly serve the people's needs for healthcare. Like health services, Red Cross practitioners are also responsible for health-related case finding, health information, provision of family planning education, home care for simple diseases, and help in organizing funerals. They serve as an additional point of entry to the healthcare system.



Source: Ministry of Health of Vietnam

Figure 4.1. Structure of the healthcare system in Vietnam

Besides this overall structure, there are 14 vertical programs of disease control and prevention in the healthcare system in Vietnam. These programs are designed for tuberculosis control, leprosy control, family planning, expanded immunization for children, malaria mitigation, control of diarrhea diseases, dengue fever, and acute respiratory infection. Each program affiliated with a central institution either under government management (MOH) or international NGOs such as UNICEF or WHO, but all of them are working through several healthcare structures. Therefore, there is no formal integration or coordination between these programs. They work in parallel with each other because they are specialized in the different specialties. Thus, there is lack of integration among them.

4.3. Health status and healthcare policy since 1975

The healthcare system in Vietnam, as elaborated in the previous section, is governed centrally and financially funded by the central government. The policy concerning healthcare is first prepared by the MOH under the supervision and guidance from the government, then passed by the parliament, and implemented by all the health care levels in the country guided by the MOH. Healthcare system and health status during the years after the country's reunification in 1975 showed comparatively good achievements, taking into account the country's backward and underdeveloped level caused by long-lasting wars and economic sanctions by the Western countries after the War with the US. However, the situation has both positively and negatively changed since the implementation of the reforms (*doi moi)* in 1986. This section attempts to look at the health status and healthcare policy before and during the comprehensive *doi moi* program conducted in Vietnam.

The structure of this section is as follows. Following this brief introduction, health status and health facilities are elaborated in section 4.3.1. Healthcare policy is discussed in section 4.3.2. Finally, a brief overview of key healthcare reforms during the *doi moi* is given in section 4.3.3.

4.3.1. Healthcare facilities and health status

Despite the heavy consequences of the war and socio-economic difficulties during the late 1970s and 1980s, the healthcare system in the country was still able to provide people with comprehensive and effective health services through many years before the *doi moi*. The system had long time been subsidized by the central government, and it functioned following the motto: "equity in healthcare and education for all people". Thus, under socialism an intensive healthcare network was built throughout Vietnam on the basic of equality in access to health services with due attention being given to primary health care in order to assure that i) all people were provided with free healthcare; ii) and people were able to access to primary healthcare services. Thanks to the government policies on free healthcare under the socialist regime, health indicators in the country were good compared to other countries in the regions as well as those with the same or higher development level in the region and in the world (see Table 4.1). In addition, while the country was still poor, backward and under-developed, in order to attain good health, the healthcare system in the country was operated and developed under the following principles: preventive medicine, people-centered[21], and cooperation and coordination between the state and the people (Do, 1995). The network of medical care has also been built in both urban and rural areas in order to create equal opportunities for all people, allowing them to be able to have access into most primary health care services and basic medication (this will be discussed more detailed in the last paragraph of this section).

With the free-healthcare mechanism during the years before the *doi moi*, the country was very successful in providing preventive health services, and control of communicable diseases through an extensively strong healthcare delivery network from the central to the primary healthcare levels including 11,000 health stations, around 1,300 inter-communal clinics, some 720 general hospitals, 80 specialized hospitals, and many other specialized

[21] People-centred approach in health care program is an approach that place people at the centre of the program. It means people are targeted for the health care programs such as diseases control and prevention medication, etc.

clinics and health centers (UN, 2003 and Vietnam Academy of Social Sciences, 2006). The healthcare network is expanding currently with the accession of a limited number of private and foreign hospitals and health centers as well as drug manufacturers and dealers through the processes of economic liberalization and equitization[22]. The numbers of doctors, health workers and practitioners, pharmacists and hospital beds per 1000 population have, therefore, increased from year to year (GSO, various years) that resulted in increase in these indicators' indices over the years (Appendix 4.3).

Besides these gains, however, there were also negative effects on health status and healthcare services caused by the economic reforms. The financial difficulties in the public and cooperative health centers caused by the transformation towards market-oriented socialist mechanism and the dismantling of the agricultural cooperatives due to market-oriented reform in the agricultural sector which had previously financially supported the basic health services (from the MOH through the hierarchical structures to the cooperative organization at the lower levels, which were governed by the level's People Committee), have weakened the healthcare system in the country. As a result, it caused non-attainment of some health plans, for example, with respect to family planning and to low rate of malnutrition among children (WHO 1978)[23]. The funding for the health sector from the government declined and became insufficient. The primary healthcare system, which used to be very stable and steady under the subsidy period, became deteriorated.

Recent statistics showed that Vietnam's public health expenditure is very low (around 1% of GDP in 2001 and around above 1.5 % of GDP in 2004) (General Statistical Office, Vietnam, various years) compared to international standards (2-5% of GDP in low-income group, 6-7% of GDP in middle-income group and more than 8% of GDP in developed countries) (United Nations, 2003). Despite this low level of public health expenditure, Vietnam's health record is good in internationally comparative terms (Table 4.1). As the data in Table 4.1 reveal, Vietnam is still very poor compared to other developing countries in the region and in the world, in terms of GNP per capita and purchasing power parity gross national income (PPP GNI) per capita. Per capita GNP (USD 550) and per capita PPP GNP (USD 2700) are just above the average of that of the low income countries (USD510 and USD 2260, respectively), and much lower than the average of the lower-middle income group (USD 1508 and USD 4630) and East Asia and Pacific region (USD 1,280 and USD 5070). But the selected health indicators are rather good in comparison to those countries and groups. Life expectancy at birth for both male and female is much higher. The infant and maternal mortality rates are much lower than the averages for the countries in the low and lower-middle income groups. In terms of country on country comparative basis, Vietnam's health indicators are favorably comparable to those of much richer countries in the upper-middle income group such as Brazil, South Africa and Turkey, as well as those of the other countries in the East Asian and Pacific region. These were results of the government policy towards providing free healthcare services under socialism.

[22] Equitization is the official English translation of the Vietnamese word co phan hoa. It is briefly and roughly defined as the process by which a part of shares of the State Own Enterprises (SOEs) are offered to the public.
[23] As pointed out by the WHO (1978), apart from political and governmental inefficiency the main reason for non-attainment of planning goals is the non-availability of adequate financial resources.

*Table 4.1. Selected economic and health indicators in selected developing countries**

Countries	GNP per capita (US$)	PPP GNP per capita (US$)	Life expectancy at birth		Infant Mortality rate (deaths per 1,000 live births)	Maternal mortality rate (per 100,000 live births)
			Male	Female		
Vietnam	550	2700	68	72	34.8	130
Indonesia	1140	3460	65	69	52.2	230
Philippine	1170	4890	68	72	36.0	200
Thailand	2540	8020	67	72	38.5	44
Brazil	3090	8020	65	73	48.1	260
China	1290	5530	69	73	30.2	56
India	620	3100	63	64	73.0	540
Peru	2360	5370	68	72	43.2	410
S. Africa	3630	10960	45	46	42.2	230
Turkey	3750	7680	66	71	48.3	70
Low income	510	2260	57	59	80.3	689
Lower-middle income	1580	4630	68	72	59.6	121
East Asia & Pacific	1280	5070	68	71	32.8	116

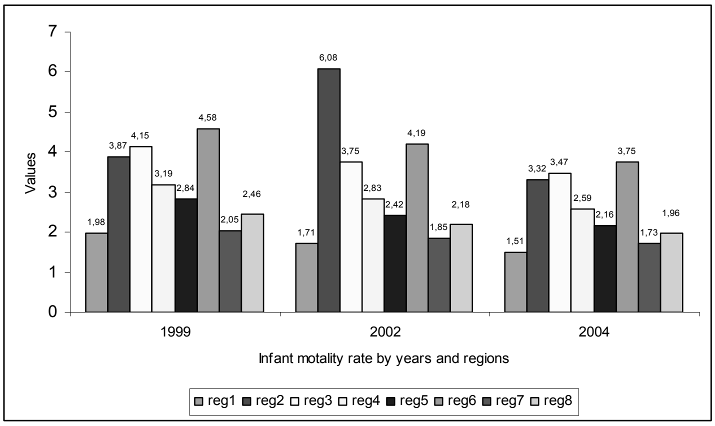
Source: World Development Report, World Bank 2006
Notes: * GNP per capita is of 2004, other indicators are of 2000-2004 (see World Development Report, World Bank (2006)

Besides those rather good records attained for the health, however, Vietnam is still facing high prevalence of malnutrition among children under five-years of age; high prevalence of low weight for children under five years of age (28.4% in 2003); high rates of maternal and neonatal mortality (130 per 100.000 cases), and high rate of induced abortions (United Nations, 2003). Also, there was a relatively high rate of people with infection of transmitted diseases such as malaria (233 cases/100,000 people), tuberculosis (TB) (116 cases/100,000 people) and HIV (110 cases/100,000 population)[24]. The high prevalence of infected diseases among people could be found in the communities, those belonging to ethnic minority, and living in remote and backward areas, where the health infrastructure and services were still poor and inadequate, education level attainment was low, and economic backgrounds were still backward. Therefore, all healthcare services were hard to reach, even at the primary health services level including preventive and promotive health services, and services for the transmissible disease control. In addition, remote and backward areas were usually characterized with bad natural conditions (such as bad weather and humidity) and inadequacy of health services and facilities together with the low level of the people's awareness towards disease prevention favored transmittable diseases such as malaria, HIV and TB, making health status in these areas lower than those in other areas and the national average.

[24] The figures are of 2004. These figures are extracted from the Vietnam Human Development Report, 2006 (Vietnam Academy of Social Sciences, 2006)

4.3.2. Spatial patterns in the distribution of healthcare facilities and health status

Generally, the health status, and healthcare facilities and services have been improved in Vietnam since the *doi moi*. However, the improvements were not distributed equally among groups and regions because the allocation of healthcare resources and budgets spent on healthcare activities (investment on health system) were not equally located among regions). The dissimilarities led to spatial differences in health status between regions and between provinces in Vietnam.



Legend:
- IMR is an acronym for infant mortality rate. 98, 02 and 04 denote the years of 1998, 2002 and 2004 respectively.
- reg1= Red River Delta; reg2 = North East, reg3 = North West; reg4 = North Central Coast; reg5 = South Central Coast; reg6 = Central Highlands; reg7 = South East; reg8 = Mekong River Delta

Figure 4.2. Spatial Evolution of Infant Mortality Rate by Regions

According to the data provided in the National Human Development Report 2006 (Vietnamese Academy of Social Sciences, 2006) health status varies sharply across regions and provinces in Vietnam. The data for the 61 provinces and by eight regions in the country show that there were spatial patterns in the distribution of healthcare facilities and health status presented by spatial differences in the selected health indicators. At the regional level, infant mortality rate (IMR) seems to be worst in the North East region which consists of the northern mountainous and backward provinces characterized by a high percentage of ethnic minorities in the population, and in the Central Highland, which is also characterized by a poor and underdeveloped public infrastructure and a high proportion of ethnic minority people. Figure 4.2 shows that the lowest infant mortality rate can be found in the Red River Delta (1.98% in 1998; 1.71% in 2002; and 1.51% in 2004) and in the South East (2.05% in 1998; 1.85% in 2002; and 1.73% in 2004). Figure 4.2 also reveals that the infant mortality rate is declining over time, from 1998 through 2004, except that there was an abnormal increase during the 1998 – 2002 period. This peak in infant mortality rate in the North East was due to an outbreak of the petechial fever in new born babies in 2002 that resulted in a large number of newborn babies that died (6.08% in 2002).

Table 4.2. Selected health indicators by regions

Regions	1998-2002-2004 LEB (year)	LEB (year)	LEB (year)	1998-2002-2004 DOC (doctor)	DOC (doctor)	DOC (doctor)	1998-2002-2004 Budg (%)	Budg (%)	Budg (%)
Red River Delta	72	73	73	42	47	50	7.7	7.6	7.2
North East	66	68	68	45	55	63	19.2	19.3	16.6
North West	65	67	67	34	43	35	23.5	20.9	23.1
N. Central Coast	70	69	70	32	43	45	11.6	11.2	11.4
S. Central Coast	71	71	72	43	49	53	10.6	9.0	9.7
Central Highlands	64	65	66	37	42	46	12.2	18.0	18.1
South East	72	72	73	38	45	48	7.6	7.3	7.6
Mekong River Delta	71	72	72	29	35	43	7.6	6.7	6.3
National average	*68*	*69*	*70*	*37.5*	*46.1*	*47.8*	-	-	-

Source: Own estimates from the original data
Legend: - All figures are rounded.
- *LEB* is the acronym for life expectancy at birth; *DOC* is number of doctors per 100,000 people; *Budg* stands for the proportion of public expenditure on health of the region to total budget expenditure. North East, North West and Central Highlands are mountainous while Red River Delta, Central Coast and Mekong River Delta are low land areas.

There were also increases in other health indicators such as life expectancy at birth, number of doctors per 100,000 inhabitants, number of nurses per 100,000 inhabitants, and budget expenditure on healthcare. As Table 4.2 reveals, the Red River Delta and the South East have the best records in the term of life expectancy at birth (72 and 73 years respectively) in comparison to other regions in the country. The statistics show that the North East has the highest number of doctors per 100,000 people (45, 55 and 63 for 1998, 2002 and 2004 respectively) and the South Central Coast is in second place with 43, 49 and 53 doctors per 100,000 population for respectively 1998, 2002 and 2004.

4.3.3. Overview of key reforms in health sector during the *doi moi*

The decline of the agricultural cooperatives and the implementation of *Khoan 10* (Contract 10) during the 1980s through the introduction of Directive 100 in agriculture[25] and the complete abolishment of cooperatives by the privatization in agriculture in the late 1980s have caused critical financing problems for the health sector. As a result, in rural Vietnam the funds from the People's Committees became insufficient and inadequate for maintaining the existing free healthcare network. Initially, the government reacted to the shortage of funds by shifting health workers at the commune level to the state's payroll in order to avoid the large scale disruption of the healthcare system during the initial period of

[25] Due to the failures of agricultural production for many years under the so-called "great socialist agriculture", the Government of Vietnam has decided to gradually decline the role of cooperative in the agriculture sector by proposing the *Khoan 10* (Contract No.10), and latter enacting the Directive 100 in the late 1980s. The contract, which aimed to lease land to households against a fixed contribution while the surplus of production was left to farmers, came into effect amidst the severe economic crisis compounded by farmers' rejection of collectivism.

reforms as experienced by the healthcare system change in China, the country having similar circumstances as Vietnam[26] (Gabriele, 2006).

The dismantling of the agricultural cooperatives and the state-lifted restrictions on private practice have encouraged the development of health services and the expansion of the network of commune health stations in Vietnam in the 1990s (Pham Manh Hung *et al.,* 2001). By implementing the economic reforms towards a market economy the government was not able to maintain the free public health system as before. The subsidy mechanism was completely abolished. Private medical practices were legitimized, the medical-drug market was deregulated, patient charges were applied in public hospitals and health centers, and the health insurance schemes were introduced into the healthcare system. Public expenditure on health has also been decentralized that allowed provincial levels to account for a large part (60%) of the total health spending.

After the legitimization of the private medical operations in 1986, the production and distribution of medical drug products were deregulated in 1989, leading to an increase in supply and availability, and a fall in the cost of drugs. By 1998, there were already more than 34,000 private medical facilities licensed. During this time, private health providers grew rapidly. As a result, a large part of health expenses shifted from the state budget to the households. As overall public health funding is short, the reform policy encourages people to use their out-of-pocket spending in order to cover the cost for public health facilities and to pay for health services received from both public health centers and private health practitioners.

The development of the private health care system with the entry of foreign investors in the health sector is consistent with the economic reform policies enacted in Vietnam. This development helps filling the gap caused by inadequacies in health infrastructure such as health clinics, drugs and facilities of the public healthcare system in the past. However, the privatization or semi-privatization in the operation of the healthcare system has caused a negative consequence as the poor tend to have little chance to access to health services, especially to those which are relatively expensive and hard-to-reach, and they therefore opt for self-medication. This option would lead to irrational over-medication and resistance to antibiotics (Segall *et al*., 2000). Another perverse consequence for the development of the country's most health facilities at the local level is that these facilities seem to be under-utilized for healthcare purposes, albeit still effective in delivering preventive health programs in the country. In his study, Gabriele (2006) found that a large share of the population in Vietnam varying from one-third to one-half, still lacks regular access to health services because they cannot afford to use the services.

In short, the increase in income among population groups and among regions brought about by the *doi moi*, and the differences in the regional policies have caused moderate differences in the health status, and health facilities and services in Vietnam (see Table 4.2 and Figure 4.2). The rapid economic growth and the process of economic reforms which induced reforms in other sectors, including the health sector, have also slightly changed the structure of the health care system in the country. Generally, the health status has been improved; healthcare facilities have also upgraded, reinforced and diversified as the statistics showed that the number of doctors, number of pharmacists and other key health indicators have significantly been improved (Table 4.2). According to the statistics

[26] In China, the disruption of the public health system caused by the disassembling of the communes was catastrophic, leading to firing of health workers (mostly in rural areas) and the collapse of the whole structure of healthcare.

provided by the GSO published in the Vietnam Human Development Report 2006, the number of healthcare centers and health workers increased both in rural and urban areas (Vietnam Academy of Social Sciences, 2006). Thus, health services, to some extent, were better able to meet the needs for people's health services. However, during the reform process the health sector was also facing many challenges. Lack of financial support to the health sector and the implementation of the privatization in the health sector led to the fact that people had to pay for their use of healthcare services and medication. This was a big challenge for poor and marginal people. This was also a disadvantage for people living in remote and backward areas because several kinds of health services were not available there, and they were also too expensive for them to have access to. Therefore, inequality in health services, healthcare facilities and health status seems to be widening among regions.

4.4. Data description and the choice of indicators

In section four we describe the data and discuss about the choice of health indicators that will be used for the analysis of inequality in the health component of welfare. However, the selection of health indicators is different among researchers. There were several groups using different approaches to measure and analyze health inequality. The first group stands for most studies in the literature on health analysis, in which each health indicator was considered as a single dimension and they were analyzed separately (see, among others, Mackenbach *et al.*, 1997; Khan, 1985). In contrast, the second group was using another approach in which the health component of welfare was analyzed under two sets of indicators namely healthcare facilities and health status using multidimensional methods (see, among others, Quadrado, 1999; Schneider *et al.,* 2002). In the third group, all the indicators associated to health were clustered together in one set, and the analysis of health was carried out for this unique set of indicators (see, among others, Ram 1982, and Morris 1979). In our study, in order to make use of the method of principal component analysis, we follow the approach used by the second group by which health was analyzed under two separate sets of indicators: healthcare facilities and health status as we argued earlier that these two sets of indicators represent two different aspects of the health component of welfare. The former stands for the investments made in the health sector, while the latter presents the outcome of such investments. In addition, the two separate analyses conducted such as: (i) an analysis of investments made in the health sector and, (ii) an analysis of health status which is considered as output or result of such investments made in (i) would help us comparing the resulting spatial distributions of the investment-based composite index and of the output-based composite index. This step would help us identifying the impact of the investments made in the health sector on the people's health status.

We organize this section as follows. A brief description of the data for analysis is given in Section 4.4.1. This section is followed by the discussion of the process of indicators selection for healthcare facilities and health status for the analysis by in section 4.4.2.

4.4.1. Data description

The empirical analysis of multidimensional health inequality and its dynamics is particularly data demanding. In general, it requires longitudinal (panel) data, which are extracted from surveys that follow each other in time and that interview the same final observational units. The data collected from the three rounds of household survey conducted in Vietnam in 1998, 2002 and 2004 (VLSS-98, VHLSS-02 and VHLSS-04) are of this kind. These surveys provide rich information on the availabilities of health services such as number of doctors, number of hospital beds, number of pharmacists and immunization for children. Information regarding health records which represent health status of people, such as life expectancy at birth and infant mortality rate, was collected from the archived sources provided by the health department at the country's provincial level and from the MOH archived at the General Statistical Office of Vietnam (GSO). Because they link the same units across time, longitudinal data are very useful for understanding the inequality dynamics, and for facilitating the measurements and assessments of multidimensional inequality in health.

The above-mentioned surveys covered 150 communes; those were randomly selected from 61 provinces in eight administrative regions (Appendix 4.1). Out of 150 surveyed communes, 120 are rural and another 30 are urban. The survey questionnaires were administered by a team of supervisors and completed with the help of local government officials and individuals in the surveyed communes. The questionnaires are filled in by direct interviews conducted by the team of interviewers trained and directed by the interview experts provided by the World Bank office and/or by the office of GSO of Vietnam. The interviewees would be the heads of households or other household members (in case of the absence of the head) who is capable to answer the questions in the questionnaires.

Other information regarding healthcare facilities and health status that is still missing from the surveys is collected from other sources such as from the MOH and/or GSO. For example, information regarding life expectancy at birth and infant mortality rate that were missing from the household surveys were gathered from the archived sources provided by the MOH and GSO of Vietnam. This information was reported to the MOH by the DOH at the provincial level.

4.4.2. Choice of indicators and their definitions

Generally, health status is commonly expressed in terms of health records such as, for example, life expectancy at birth and infant mortality rate. Health facilities refer to infrastructure for the healthcare system. Healthcare infrastructure in a territory constitutes health services and - facilities available to the use of people living in such a territory. Health infrastructure represents the number, location, size of healthcare facilities and manpower relative to a potential user population as well as the health services provided to people for disease treatment and prevention. In other words, for the case of the study of health inequality the data should cover both the health status and the availability of the health services in a territory chosen as a spatial unit for analysis.

Several issues related to the selection of specific indicators for the health component within this study are worth to be mentioned and discussed. Firstly, the selected indicators must well capture a majority of dimensions in health status, health facilities and services for which data are available. So they could adequately reflect the actual health status of the people and health services available to people in a geographical territory or unit (e.g. province in our study). Secondly, the selected indicators should well be presented and able to be primarily intended to support regional health authorities in monitoring the progress in improving and maintaining the health of the population and the functioning of the health system.

On the basis of the above considerations and the availability of the relevant data, the following two sets of indicators were selected for the analysis.

a. Indicators for healthcare facilities

1. Childhood immunization*:* the proportion of children who have been fully immunized against diphtheria, pertussis, tetanus, Haemophiliac influenza type b (Hib), measles, numbs and rubella;
2. Hospital beds*:* the number of hospital beds per 100,000 people;
3. Medical doctors: the number of active registered civilian general practitioners per 100,000 people;
4. Pharmacists: the number of active registered drug chemists per 100,000 people.

b. Indicators for health status

1. Life expectancy: an estimate of the average number of years that a person born in a certain year is expected to live based on current mortality rate;
2. Infant mortality: the number of infants, who die in the first year of life, is expressed as a rate per 1,000 live births.

These two sets of indicators will be used for the analyses of the two dimensions of health separately. The first dimension, which contains the variables: hospital beds, childhood immunization, medical doctors and pharmacists, stands for the healthcare facilities available to the use of people. These variables were considered as the investment made in the health sector by both the local and central governments of Vietnam, and also by private spending on health. The second dimension, which includes the variables life expectancy and infant mortality, partly represents the health status of people. This dimension reflects the returns to the investment made in the health sector.

4.5. Spatial interrelation and inclusion of spatial contiguity in health facilities:

In section 4.5, we discuss the spatial interrelation and the procedure of the inclusion of spatial contiguity in health facilities for the analysis. We structure this section as follows. The discussion of the spatial spillover effects in health is given in section 4.5.1. In section 4.5.2, we develop the framework to incorporate the spatial contiguity into the analysis of spatial inequality in health facilities.

4.5.1 Spatial spillover effects in health facilities

People living in one province are able to access to healthcare facilities available in adjacent province(s). Spillover effects in health facilities are therefore in existence and they should be taken into account in order to examine the spatial effects of the healthcare facilities among provinces and the effects of the healthcare facilities in one province on health status in adjacent provinces. For the case of health status, we argue that the health status of people living in one region would not have any impact on the health status of people living in adjacent regions. Thus, we don't take contiguity effects into consideration in the analysis of spatial inequality in health status.

Spillover effects or spatial effects are a result of the contiguity relationship. The contiguity reflects the relative position in space of one regional unit of observation to another one. For the spillover effects, we consider only the first-order contiguity because it is argued that people usually go to the closest regions for medical treatment. In addition, using the first-order contiguity would help to simplify the process of spatial inequality analysis. In practice there were cases in which people seek for the higher quality treatment by medical experts that was unavailable neither in their province nor in adjacent provinces in its first-order contiguity. In another case, people may choose to seek for treatment in other provinces rather than the first-order ones, where the service for the treatment they need is available. However, these possibilities are there but very rare in practical life. Thus, we therefore omit these possibilities because they would complicate the study. Instead, we focus on the use of the first-order contiguity only.

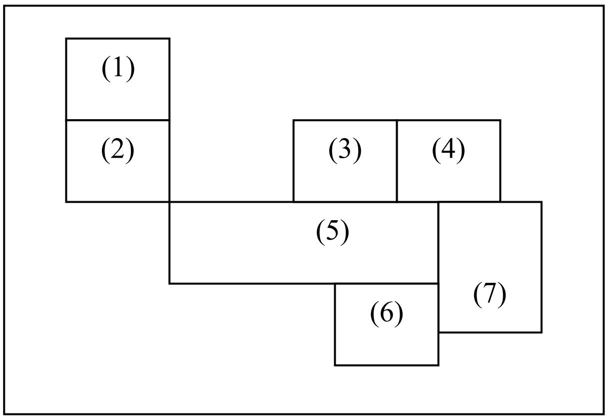


Figure 4.3. An illustration of contiguity

For the definition of contiguity, Hordijk (1974), Folmer (1986) and LeSage (1999) defined the neighboring or contiguous provinces according to the notion of order of contiguity in space. As illustrated in Figure 4.3 that shows a hypothetical example of seven regions as they would appear on a map. The first order of contiguity implies that two geographical or spatial units (e.g. regions, provinces) share a common boundary of a non-zero length as region (1) and region (2) (*rook contiguity*), or share a common vertex as region (2) and region (5) (*bishop contiguity*), or they share both a common boundary and a common vertex as region (5), region (7) and region (2) (*queen contiguity*). A second order of contiguity is defined between two non-contiguous spatial units that are being separated by one unit as region (6) and region (3) which are separated by region (5).

4.5.2. Inclusion of spatial contiguity in healthcare facilities

In this section we elaborate the procedure to incorporate contiguity into the analysis of healthcare facilities. This procedure is applied in this chapter for the analysis of spatial inequality in healthcare facilities only. Spillovers are drawn from comparing the findings from two cases: contiguity and non-contiguity of the health facilities.

As we argued earlier, the resources for use in one province are determined by incorporating the resources available in this province and its neighboring provinces. According to Nijkamp (1978) and Quadrado (1999), the level of available resources for use (A_i) in province *i* is the sum of the resources available of its own (of *i-th* province) (A_i^o) and those available in the adjacent provinces (A_i^a). It is presented in the following equation:

$$A_i = A_i^o + A_i^a \qquad (4.1)$$

As mentioned above, A_i^a is the resources in the adjacent provinces to province *i* that could be accessed by people living in province *i*, so it is defined by using spatial weight as follows.

$$A_i^a = \sum_{j=1}^{n} w_j A_j \qquad (4.2)$$

where: A_j is the resources available in province *j* that are adjacent to province *i*. *j=1,...,n* (*n* is the number of provinces that are adjacent to province *i*), and w_j is the spatial weight.

Anselin (1988) suggested that spatial weights could be expressed based on a function of inverse distance known as 'the simple inverse distance or any integer power of the inverse distance'. Thus, the distance between two adjacent provinces (exactly the distance between the two provincial capitals) measured by travel inputs (say travel time or travel distance) is raised to one or other integer powers, say *p*. So the spatial weight is defined as follows.

$$w_j = \frac{1}{d_{ij}} \qquad (4.3)$$

where d_{ij} is the distance between the two provincial capitals of the province *i* and its adjacent province *j*.

From equation (4.3) we observe that the spatial weight is defined as the simple inverse distance between two provincial capitals of the adjacent provinces[27]. For the case of healthcare facilities, the distance between two adjacent provinces is considered to be very important in the people's decision to look for medical treatment, so it is worthy to incorporate it into the analysis.

[27] Another alternative way to compute the weight is the simple inverse of the average traveling time between two capitals of the two adjacent provinces, i.e. the denominator of equation 4.3 is replaced by the traveling time that it takes to go from one provincial capital to the capital of the adjacent province.

4.6. Results

As discussed earlier, results will be reported for healthcare facilities and health status separately. For the case of healthcare facilities, analysis will be carried out for both the non-contiguity and the contiguity case in order to explore the dynamics of regional inequality and to evaluate the importance of spillover that occurred among regions/provinces because of co-sharing of health facilities between adjacent provinces. The organization of this section is as follows. Measurement and analysis of multidimensional regional inequality in healthcare facilities is provided in section 4.6.1. The results for the health status are given in section 4.6.2.

4.6.1. Healthcare facilities

We structure this section as follows. The results for the healthcare facilities in the case of non-contiguity are given in section 4.6.1.1. Section 4.6.1.2 presents the results for the healthcare facilities in the case of contiguity.

4.6.1.1. Healthcare facilities: the non-contiguity case

In the case of non-contiguity, we assume each province to be a single and independent entity. Thus, all health facilities available in one province serve only residents of that province.

We organize the section as follows. The analysis of PCPC is given in section 4.6.1.1.1. This is followed by the measurement and analysis of regional inequality in section 4.6.1.1.2.

4.6.1.1.1. The analysis of the partial Common Principal Component (PCPC)

We follow step 1 to step 3 presented in section 3.5 of chapter three to estimate the PCPC model for all three periods and to test the null hypothesis that there is a common component for all three periods against the alternative that the components of these periods are not the same. The test results provided in Table 4.3 show the value of the likelihood ratio statistic: χ^2 = 1.704 on 6 degrees of freedom with a very high p-value (0.945). The computed Chi square (1.704) is very well below the critical χ^2 at 95% significant level on 10 degrees of freedom (critical χ^2=12.95). With such a high p-value and very low computed χ^2 compared to the critical value, we are unable to reject the null hypothesis at almost all significant levels. Therefore, we strongly accept the null hypothesis and confidently conclude that there is a common principal component for all the periods under consideration. These test results will be further supported by checking the proportion of variance recovered by the first common component, standard errors of the estimated coefficients, and the matrices of covariance and correlation between the estimated principal components. As proven by Flury (1988), apart from the acceptance of the null hypothesis the model fits the data well when: i) the proportion of variance recovered by the first common principal components should be reasonable (at least 60% variance recovered from the original variables by the first common principal component); ii) the standard errors of the common component coefficients are not larger than 0.1; and iii) under the null hypothesis of the PCPC, correlations between estimated components are close to zero.

Table 4.3. Partial common principal component test

Number of estimated parameters in the model	24
Likelihood ratio test χ^2	1.704
Number of degrees of freedom	6
P-value	0.945
Critical χ^2 at 95% significant level on 6 degrees of freedom	12.59

Table 4.4 displays the maximum likelihood PCPC estimates of the common component coefficients for all the three periods with corresponding standard errors[28], and the proportion of variance recovered by the first common principal component. Table 4.4 shows all coefficients have expected (positive) sign. Their corresponding standard errors are reasonable (<0.1). In addition, the variance recovered from the total variance of the original variables by the first common component was very good (79 % for 1998; 87 % for 2002 and 80 % for 2004). The discarded components would lose only 21 %, 13 % and 20 % of total variance for the 1998, 2002 and 2004 respectively. According to Flury (1988), the proportion of variance recovered from the total variance by the first common principal component was very high in all years. From the statistical point of view, it is therefore reasonable to conclude that the indicators' coefficients of the first common principal component can be considered as robust and reliable for interpretation.

Table 4.4. Maximum likelihood PCPC estimates of eigenvectors and characteristic roots†

Eigenvectors (common component coefficients)			
Childhood immunization		0.178	(0.022)
Hospital beds		0.291	(0.079)
Medical doctors		0.379	(0.090)
Pharmacist		0.710	(0.098)
Characteristic roots of the first common component			
	1998	2002	2004
Characteristic roots	0.049	0.039	0.040
Proportion of total variance recovered	79 %	87 %	80 %

† Standard errors are in parentheses

Table 4.5 displayed the covariance and correlation matrices between the estimated components for all periods based on the estimated PCPC model. We observe from this table that there were no substantial correlations between the common principal components (except some moderate correlations of 0.086 between the first and the fourth principal component in the matrix for the 1998 sample, and 0.132 between the first and the third principal component in the matrix for the 2002 sample) occur in all the three years. The correlation coefficients were very small in almost all cases. These small correlations therefore further support our conclusion about the acceptance of the null hypothesis of the common principal components for the three years. Together with the maximum likelihood test and the checking of the proportion of variance recovered by the first common principal component we finally conclude from a statistical point of view, that the coefficients from the first common principal component are robust and reliable. These coefficients can be

[28] The standard error of the coefficients (first common principal components) estimated from PCPC model is computed using equation (3.63). (See appendix 3.3 in chapter three)

used for interpretation and analysis of the composite index and the regional inequality with respect to healthcare facilities.

Looking at the estimates displayed in Table 4.4, the indicators' coefficients of the first common component, we find that the variables pharmacist and medical doctors are the most dominant variables in the component of healthcare facilities (their coefficients values are 0.710 and 0.379 respectively). It means that these variables have the highest impact on the values of the composite index, and then on regional inequality in health. In contrast, variables such as childhood immunization and life expectancy are of less importance as their coefficient values are smaller than those of other indicators (0.178 and 0.291, respectively). This result is in conformity with the fact that the gap in life expectancy at birth between regions is practically quite small in Vietnam, and that the immunization of children is a nation-wide program implemented by the MOH under the guidance from the government. The government sets a goal to sufficiently provide a full immunization scheme to all children irrespective of region, social background and ethnic group. Therefore, all the children were equal in accessing to immunization. Thus, the gap in the rate of children who are fully immunized between regions in the country is quite small.

Table 4.5. Covariance (F_i)[29] and correlation (R_i)[30] matrices of the partial PCs for 1998, 2002 and 2004[@]

Matrix for the 1998 Period				
	1st PC	2nd PC	3rd PC	4th PC
	0.049	0.012	-0.014	0.008
	0.075	0.007	0	0
R_{98}/F_{98} =	-0.026	0	0.003	0
	0.086	0	0	0.003
Matrix for the 2002 Period				
	1st PC	2nd PC	3rd PC	4th PC
	0.039	0.009	0.017	0.014
	0.053	0.001	0	0
R_{02}/F_{02} =	0.132	0	0.002	0
	-0.076	0	0	0.004
Matrix for the 2004 Period				
	1st PC	2nd PC	3rd PC	4th PC
	0.040	0.008	0.015	0.0019
	0.042	0.005	0	0
R_{04}/F_{04} =	0.067	0	0.004	0
	0.093	0	0	0.001

[@] On and above diagonal of the matrix - variances and covariance; below the matrix diagonal - correlations

---

[29] As in an ordinary PCA analysis, we can compute covariance and correlation matrices between estimated principal components for the case of PCPC analysis. In the present case, the lower right (p-q)×(p-q) portion of the covariance matrices is calculated using formula: $F_i = \hat{\beta}^{(i)'} S_i \hat{\beta}^{(i)}$, where $\hat{\beta}^{(i)}$ is consistent estimate of $\beta^{(i)}$ and $\beta^{(i)'}$ is transposed of $\beta^{(i)}$. S_i is covariance matrix of indicators.

[30] Similarly, the formula to compute the correlation matrices can be written as: $R_i = \hat{\Lambda}_i^{-1/2} F_i \hat{\Lambda}_i^{-1/2}$, where $\hat{\Lambda}_i$ is a consistent estimate of Λ_i, $\Lambda_i = diag(\lambda_{i1}, \ldots\ldots, \lambda_{ip})$, while F_i is the covariance matrix computed using the formula given in note 10.
For more details about these formulas and how to construct the matrices of variances and correlations, we refer to the book by Flury (1988), p. 71 and p. 131, and the relevant examples demonstrated there.

4.6.1.1.2. The analysis of multidimensional regional inequality for healthcare facilities

We follow step 9 presented in section 3.5 of chapter three to determine the weights associated to the indicators using the estimates of the coefficients provided in Table 4.4 and then to compute the composite index and Theil's second measure. The procedure followed for obtaining weights is as follows. The characteristic vector corresponding to the largest characteristic root of the covariance matrix of the four constituents (indicators/variables), derived from the estimates of PCPC model, is rescaled to make the sum of its elements equal to one, and the common principal component is a weighted sum of the four constituents with rescaled elements of the characteristic vector as the weights (Morris, 1978; Ram, 1982). These weights help to determine the relative importance of each element on the indicators' coefficients. We formularize this procedure in:

$$w_{si} = s_i \Big/ \sum_{i=1}^{n} s_i \qquad (4.4)$$

where w_{si} is the weight attached to the indicator *i*, s_i is the coefficient of indicator *i* obtained by either PCPC or PCA, and *n* is the number of indicators. Applying the above formula to the estimated coefficients given in Table 4.4, the associated weight for each indicator, therefore, is easy to be derived and displayed in Table 4.6.

Table 4.6. Weights associated to the indicators in the common component

Indicators	Weights
Childhood immunization	0.114
Hospital beds	0.187
Medical doctors	0.243
Pharmacist	0.458

Using the weights given in Table 4.6 and the indicators' index provided in Appendix 4.3, composite indices[31] for each province were computed. The composite indices and ranking order by provinces are given in Table 4.13.

Table 4.7. Theil's second measure for the health component and its decomposition

	1998		2002		2004	
	Value	Share	Value	Share	Value	Share
Between-Region inequality	0.075	44%	0.063	43%	0.072	46%
Within-Region inequality	0.094	56%	0.083	57%	0.086	54%
Total inequality	0.169		0.146		0.158	

Theil's second measures have been computed for all three years and are displayed in Table 4.7. We can observe from this table that total inequality has expressed the up and down movements during the entire period from 1998 to 2004 as the total index (total inequality) decreased during the 1998-2002 period (from 0.169 in 1998 down to 0.146 in

[31] Because all the indicators with respect to health status and healthcare facilities are chosen and transformed to be the positive indicators, thus, we come to terms that the higher value of the composite index means a better position and more favorable status for the corresponding region/province in term of the health component of welfare.

2002), and again increased during the 2002-2004 period (from 0.146 in 2002 up to 0.158 in 2004). However, total inequality generally decreased during the period. The decrease means that the gaps in healthcare facilities between regions were getting closer. The decrease is in contrast to the theory which says that the inequality in a country is increasing at its initial stage of development. This can be explained as that during the *doi moi* process, the health sector of the country was partly privatized because the subsidy system in the health sector in Vietnam has been broken down due to socioeconomic reforms. As a result, more investment from both the foreign and domestic investors accrued to the health sector. Roughly estimated, there were a dozen projects with total capital amounting to hundred million dollars that have been invested in the health sector. Therefore, healthcare service was better equipped and well functioning. Generally, in almost all regions in the country people got more chances to get access to up-to-date healthcare facilities and good medicines. All these events have made the regional inequality in healthcare facilities to decline in the 1998-2004 period.

The decomposed components of the total inequality as between-group and average within-group inequality are also reported in Table 4.7. These decomposed terms are very informative. They can help us seeing whether the major changes in healthcare facilities were caused by changes within groups or between groups. Table 4.7 shows the different changing patterns between the two terms over the times. The between-region inequality decreased during the 1998-2002 period from 0.075 to 0.063, and then increased in the next period, from 0.063 (in 2002) to 0.072 (in 2004). The changes in the between-region inequality were still very moderate. In a similar fashion, the average within-region inequality has also changed in the 1998-2002 period, from 0.094 in 1998 to 0.083 in 2002, and then lightly increased in the 2002-2004 period, from 0.083 to 0.086. As we observe in Table 4.7, the within-region inequality makes up a larger fraction (56 %, 57 % and 54 % for the 1998, 2002 and 2004 respectively) than the between-region inequality does (44 %, 43 % and 46 % for the respective 1998, 2002 and 2004) in the overall inequality in all the three years under study. This finding can be explained as follows. As the government policy towards equality in health irrespective of region, social background and ethnic group has played an important role in making between-region inequality more moderate. However, this is macro policy, so it impacts at macro levels. It means, that the distribution of health facilities within the region is, to some extent, out of reach of such a policy. So, it makes the within-inequality larger. In addition, another reason for a moderate level of between-region inequality was that resources of health facilities and services were rather equally allocated among regions following the government policy towards equality in health between regions in the country.

4.6.1.2. Healthcare facilities: the contiguity case

We organize this section as follows. The analysis of PCPC is given in section 6.1.2.1 and followed by the analysis of regional inequality in section 4.6.1.2.2. The analysis of the impact of spillover effects is provided in section 4.6.1.2.3.

4.6.1.2.1. The analysis of Partial Common Principal Component (PCPC)

We follow step 1 to step 3 in section 3.5 of chapter three to estimate the PCPC model for the three periods under study in order to test the null hypothesis which states that all the periods share a common principal component (H_{pcpc}) versus the alternative that each period has a distinct component (H_{pca}). The results of the test are given in Table 4.8. According to the test results, the value of the likelihood ratio statistics: χ^2 = 1.42 at 6 degrees of freedom (critical χ^2 at the 97.5 significant level with 6 degrees of freedom is 14.45), and p-value = 0.96. With this high p-value and the calculated χ^2 very well below the critical value, we are very confident to accept the null hypothesis of the common principal component for the three samples.

Table 4.8. Partial common principal component test: the case of contiguity

Number of estimated parameters in the model	24
Likelihood ratio test χ^2	1.42
Number of degrees of freedom	6
P-value	0.964
Critical χ^2 at 97.5% significant level at 6 degrees of freedom	14.45

We go on to take step 4 in section 3.5 of chapter three. Table 4.9 shows the exact minimum likelihood estimates of the common component coefficients for the three years under study including their corresponding standard errors and the proportion of total variance recovered by the first common principal component for each period. We observe from this table that all the estimated coefficients have the expected sign and reasonable standard errors. The proportion of variance recovered from the total of variance by the first common component was very high (80 % for the 1998 sample, 82 % for the 2002 sample and 80 % for the 2004 sample), which implies that the model fits the data very well and the estimated coefficients are robust from a statistical point of view.

Table 4.9. Maximum likelihood PCPC estimates of eigenvectors and characteristic roots (Contiguity case)†

Eigenvectors (common component coefficients)			
Childhood immunization		0.189	(0.036)
Hospital beds		0.299	(0.043)
Medical doctors		0.411	(0.068)
Pharmacists		0.785	(0.096)
Characteristic roots of the first common component			
	1998	2002	2004
Characteristic roots	0.05	0.041	0.041
Proportion of total variance recovered	80 %	82 %	80 %

† Standard errors are in parentheses.

The matrices of the covariance and correlation between estimated components for the case of contiguity are given in Table 4.10. The table shows that, like in the case of non-contiguity, there were no substantial correlations between the estimated components for all the periods except that moderate correlations occurred between the first and fourth component in the 1998 sample. These findings further consolidate our acceptance of the null hypothesis and our conclusion about the robustness of the first common component's coefficients.

Table 4.9 also reveals that all the indicators coefficients are of the expected sign. Like the case of non-contiguity, the variables pharmacist and medical doctors were still the most dominant, while the least dominant variable was the childhood immunization.

Table 4.10. Covariance (F_i)[32] and correlation (R_i)[33] matrices of the partial CPCs for 1998, 2002 and 2004@: the case of contiguity

	1stPC	2ndPC	3rdPC	4thPC
	0.050	-0.015	0.0085	0.0034
R_{98}/F_{98} =	-0.0682	0.007	0	0
	0.0519	0	0.003	0
	0.1125	0	0	0.003
2. Matrix for the year 2002				
	1stPC	2ndPC	3rdPC	4thPC
	0.041	0.0085	-0.0124	0.0014
R_{02}/F_{02} =	0.0684	0.003	0	0
	-0.0568	0	0.002	0
	0.0576	0	0	0.004
3. Matrix for the year 2004				
	1stPC	2ndPC	3rdPC	4thPC
	0.040	0.0076	0.0118	0.0025
R_{04}/F_{04} =	0.0329	0.004	0	0
	0.108	0	0.004	0
	0.0823	0	0	0.002

@ On and above diagonal of the matrices - variances and covariances; below the matrices' diagonal - correlations.

[32] See Note #29
[33] See Note # 30

4.6.1.2.2. The analysis of multidimensional regional inequality for healthcare facilities

Using equation (4.4), the weights[34] attached to the indicators of the healthcare facilities component derived from the indicators' coefficients provided in Table 4.9 are displayed in Table 4.11. The weights are used to compute the composite indices for the healthcare facilities in the case of contiguity, and then the Theil's second measure of inequality in health facilities and its decomposed components, see Table 4.12.

Table 4.11. Weights attached to indicators in the first common component: the case of contiguity

Indicators	Weights
Childhood immunization	0.112
Hospital beds	0.178
Medical doctors	0.244
Pharmacist	0.466

Table 4.12 Theil's second measures and its decomposition: the case of contiguity

Periods	1998		2002		2004	
	Value	Share	Value	Share	Value	Share
Between-region inequality	0.072	53 %	0.067	49 %	0.069	51 %
Within-region inequality	0.064	47 %	0.071	51 %	0.066	49 %
Total inequality	0.136		0.138		0.135	

As Table 4.12 reveals, the total inequality remained more or less stable. In contrast, the patterns of the between-region and within-region inequality were different. The between-region inequality declined from 0.072 in 1998 to 0.069 in 2004, while the within term went moderately up from 1998 to 2002, from 0.064 to 0.066. In the case of contiguity, the share of between region and within region inequality in the total inequality was getting closer than in the case of non-contiguity. Total inequality remained more or less stable.

4.6.1.2.3. The impact of the spillover effects

When comparing Tables 4.7 and 4.12, we observe that total inequality in the non-contiguity case was higher than that in the contiguity case in all the periods. However, if we take a closer look at the decomposed terms of inequality we found that the share of each term has changed in the case of contiguity compared to the case of non-contiguity. In the case of contiguity the share of within-region inequality is less than that of between-region inequality except for 2002, while in the case of non-contiguity the former makes up a larger part than the latter. These changes are further presented in Figures 4.3, 4.4 and 4.5.

Figure 4.4, the trends of total inequality: contiguity versus non-contiguity shows that in both cases the total inequality has declined. However, it was up first then down later in the former case, while it decreased first then increased later in the latter case. As the figure shows, the change in total inequality in the case of contiguity was smoother than the change in the case of non-contiguity. For the trends of between-region inequality, figure 4.5

[34]For the procedure to follow for obtaining these weights, see section 5.1.1.2 in this chapter

tells us that in both cases the patterns of the between inequality were the same; decrease then increase. Finally, the within region inequality was in complete contrast between the contiguity and non-contiguity cases (see figure 4.6).

These results indicate the importance of spillover effects in regional inequality in healthcare facilities. It suggests that it is necessary to account for the spillover in the measurement and analysis of regional inequality in healthcare facilities in Vietnam. It is likely that the non-contiguity case is to give biased results, because in this case we do not take into account the fact that people are able to have access to the health facilities and services available in neighboring provinces in its first-order contiguity relation.

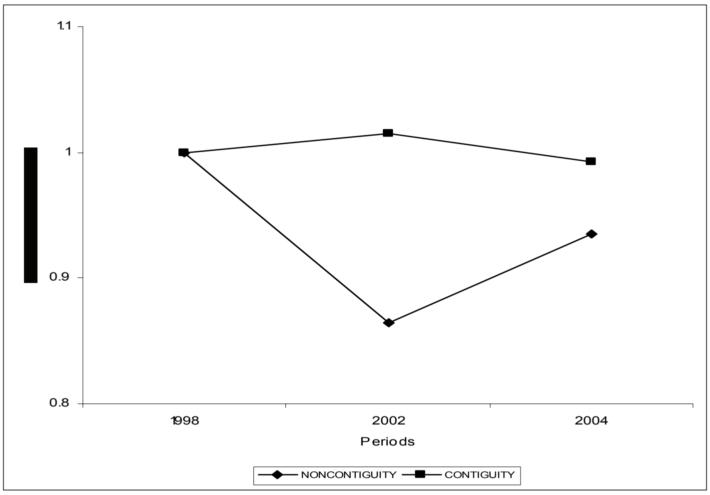


Figure 4.4. Trends of total inequality: contiguity versus non-contiguity

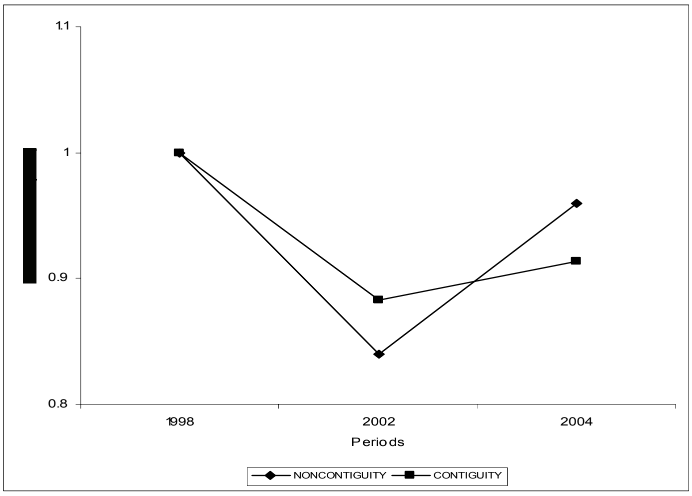


Figure 4.5. Trends of between-region inequality: contiguity versus non-contiguity

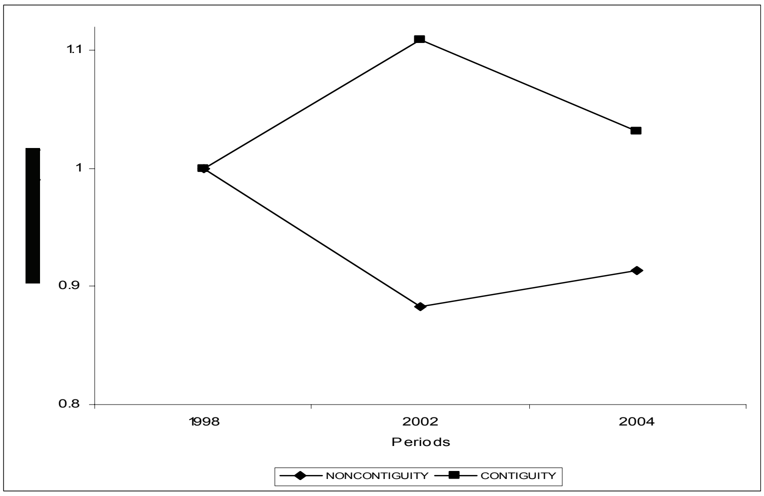


Figure 4.6. Trends of within-region inequality: contiguity versus non-contiguity

Note: Theil's second measure of 1998 is equated to one and the indices of the 2002 and 2004 are adjusted proportionally.

The composite indices in 61 provinces for both the contiguity and non-contiguity cases[35] were computed and are reported in Table 4.13 and graphically presented in Figures 4.7, 4.8 and 4.9 for respectively 1998, 2002 and 2004 for both the contiguity and non-contiguity cases. These figures show the clusters of most-favored, medium-favored and least-favored regions with respect to healthcare facilities. The regions with the higher values of composite indices are considered as more favored than the ones with lower values of the composite indices.

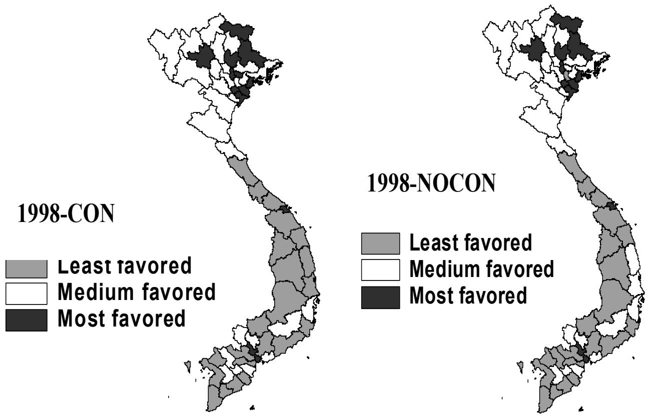


Figure 4.7. Clusters of favored regions with respect to healthcare facilities for the 1998 sample: contiguity (CON) versus non-contiguity (NOCON)

[35] The least-favored regions with respect to healthcare facilities are those with the value of the composite index less than 0.4, for medium-favored region, the composite index ranges from 0.4 to 0.55, and for the most-favored region the composite index is larger than 0.55 for all the periods and for both the contiguity and non-contiguity cases.

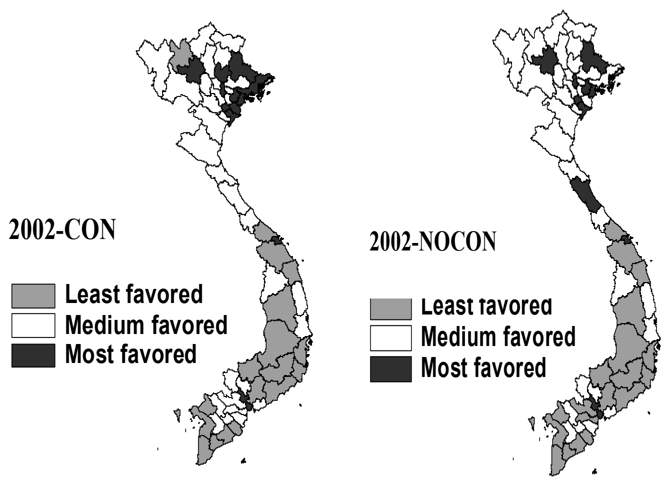


Figure 4.8. Clusters of favored regions with respect to healthcare facilities for the 2002 sample: contiguity (CON) versus non-contiguity (NOCON)

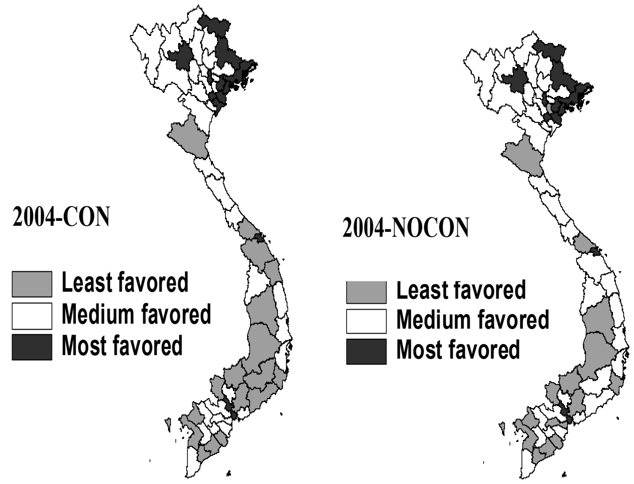


Figure 4.9. Clusters of favored regions with respect to healthcare facilities for the 2004 sample: contiguity (CON) versus non-contiguity (NOCON)

The Figures 4.7, 4.8, and 4.9 present the clusters of favored regions over time for two cases: contiguity versus non-contiguity. The figures reveal that the Red River Delta, some parts of the North East and the South East (Ho Chi Minh City) were the most-favored regions in both the contiguity and the non-contiguity case. The most-favored regions include the lowland provinces with a higher level of development such as Hanoi, Hai Phong, Quang Ninh in the North, and Ho Chi Minh City and Binh Duong in the South. This means that the most-favored regions with respect to healthcare facilities did not change over time in both the case of contiguity and of non-contiguity. Meanwhile, medium-favored regions changed over periods for the case of contiguity compared to the non-contiguity one. The medium-favored regions comprised most parts of the North West, the

North Central Coast, parts of the South Central Coast and some parts of the Mekong River Delta. The least-favored regions were those characterized by a low level of development in the South Central Coast, the Central Highlands and some parts of the Mekong River Delta, in which the level of public spending on health were lower than in other regions and even smaller than the national average (see Table 4.2).

Finally, taking a closer look at the composite indices of 61 provinces as displayed in Table 4.19, we find that the Red River Delta records the highest level of the composite index in all three years, which implies that the region is most-favored with respect to healthcare facilities. The Central Highland was registered with the lowest composite index which means that the region is least-favored in term of healthcare facilities despite the fact that this region has received the highest share of budget recently. This can be explained as that, in the past the health facilities in the region were very poor and backward and in severe shortage. Recently high investment in the health sector accrued to this region just partly to compensate for such a shortage and to upgrade currently available health facilities of the region. Such investment could not be able to make the health facilities of the region better than the one of other regions with good health facilities such as the Red River Delta. Table 4.19 also tells us that the differences in composite indices between the samples are reasonable. If we look at the ranking order by regions, there is evidence showing the changing patterns in health indicators between the surveys. We observe that the ranking order of most regions do not change in the 1998-2006 period except that there are swaps between the North East and the North West, and between the South East and the South Central Coast.

The high level of the composite indices scored by the Red River Delta and the South East can be explained. Firstly, these two regions are located with the most developed and wealthier cities and provinces such as Hanoi, Hai Phong, Hai Duong, Nam Dinh and Thai Binh in the former case and Ho Chi Minh City, Ba Ria-Vung Tau, Binh Duong and Dong Nai in the latter, which rank higher than other regions. In these two regions many economic zones, export processing zones and international offices as well as offshore and foreign companies are to be found, which attract many highly educated. Therefore, these regions are characterized by a relatively higher number of medical doctors per 100 thousand populations (42 for the Red River Delta and 38 for the South East) than the national averages (37), higher life expectancy at birth (73.2 for the former and 71.6 for the latter) than the national average (at 69.8) for the 1998 sample. Similar results are found for the 2002 and 2004 samples as all the health indicators were higher than the corresponding national averages (see Table 4.2).

Meanwhile, the least-favored regions such as the Central Highlands are characterized by poor, backward and mountainous provinces. A large part of the people living in these provinces is of several ethnic minority groups such as Ede, Gia-rai and H'mong. These regions, which rank in terms of the composite indices lower than others, are least developed in the infrastructure. This is an exceptional case because there are inadequate healthcare facilities such as hospitals, number of doctors, pharmaceutical drugs, disease treatment and preventive measures despite the high rate of budget spent (18%) on health for this region (Table 4.2). The reason is that despite the fact that more money was spent on the health sector for the Central Highlands, there is still to be found less development and a shortage of health infrastructure in this region. These shortages make health indicators less advantageous than the national averages. In addition, the low levels of education, the backward and unhealthy lifestyles have also caused increasing infectious diseases amongst people. As statistics show, the number of doctors per 100,000 people in the Central Highlands was 37, while the national average stood at 37 for 1998. These figures have increased to 42 and 45 respectively in 2002 and 46 and 49 in 2004. For other health indicators under consideration, the same trend is found. All coefficients of the indicators in the least favored regions/provinces are lower than those of the most favored regions/provinces and also lower than the national averages (GSO, various years).

The calculated composite indices by regions and provinces also provide justification for the initial observation that the changes in spatial inequality in healthcare facilities are rather moderate as all the indices show small differences between the three years under analysis. The swaps in ranking among regions are not so common, which implies that the regions, that were good in health status and health facilities, are still favorable with respect to healthcare facilities.

4.6.2. Health status

In section 4.6.2 we examine and analyze differences in health status among regions in Vietnam. The data used for this analysis is the same as the one used for the analysis of spatial inequality in healthcare facilities (see the data description in section 4.4.2). The selected indicators and their definition were already discussed in section 4.4.1. As argued earlier we will not consider contiguity effects for the health status because health status of one region seemingly has no effect on the health status in other regions.

This section is structured as follows. The analysis of partial principal component model and the results of the composite index estimation for health status are given in section 4.6.2.1. Then, the analysis of spatial inequality for health status is provided in section 4.6.2.2. In section 4.6.2.3, we will compare spatial patterns between healthcare facilities and health status.

4.6.2.1. The analysis of the Partial Common Principal Component (PCPC)

We follow step 1 through step 3 presented in section 3.5 of chapter three to estimate the PCPC model for all three periods in order to test the null hypothesis of a common principal component among three periods under study versus the alternative that each period has its own component. The test results are displayed in Table 4.13. The test results tell us that we cannot reject the null hypothesis even 10% of significant level because of the low value of the computed χ^2 (1.96) and such a high p-value (0.375).

Table 4.13. Partial Common Principal Component test for health status

	PCPC(1) for all samples
Estimated parameters	7
Likelihood ratio test χ^2	1.96
Degrees of freedom	2
P-value	0.375
Critical χ^2 at 10 % of significant level at 2 degree of freedom	4.61

Table 4.14. Eigenvectors and Maximum Likelihood PCPC estimates for health status

Variables	Eigenvectors (common component coefficients*)		
Life expectancy	0.853 (0.098)		
Infant mortality	0.425 (0.053)		
Characteristic root of 1st PC	1998	2002	2004
Characteristic roots	0.0016	0.0013	0.0012
Proportion of total variance	0.98	0.99	0.99

Standard errors in parenthesis.

From Table 4.14, which contains the principal component coefficients, their standard errors and maximum likelihood PCPC estimates for all the periods, we find that all the coefficients have reasonable standard errors. In addition, the variance recovered by the first common principal component was very high in all the cases (close to 100%). With the likelihood ratio test results above, we finally conclude that the PCPC model fits the data very well. Therefore, all the estimated coefficients (the first common principal components) are robust and reliable. Thus they can be used for the analysis of the composite index with respect to health status, and then in the analysis of spatial regional inequality.

The estimated coefficients presented in Table 4.14 tell us that the variable life expectancy was the dominant indicator in the composite index with respect to health status while infant mortality is less dominant. However, the value of both coefficients is quite high. This means that the differences in effects that the two variables have on the composite index and then on the regional inequality are not so large. The variable life expectancy is more dominant because the life expectancy at birth was a bit different among regions while infant mortality was closer among regions. We find high life expectancy at birth in the lowland areas and low life expectancy at birth in the mountainous areas (see Table 4.2).

4.6.2.2. The analysis of multidimensional regional inequality for health status

We take Step 9 (section 3.5 of chapter three) to determine the weights[36] associated to the indicator's coefficients presented in Table 4.14, and then to compute composite index and Theil's second measure for the health status. The weights are displayed in Table 4.15,

[36] For the estimation of these weights, we refer to section 4.5.1.1.2.

the composite index of 61 provinces with respect to health status is provided in Table 4.20, Theil's second measure and its decompositions are given in Table 4.16.

Table 4.15. Weights attached to indicators in the first common component for health status

Indicators	Weights
Life expectancy	0.667
Infant mortality	0.333

The decline of Theil's second index for the health status (see Table 4.16) tells us that total inequality steadily declined over time. This means that the health status was improving across population and regions in Vietnam. It is interesting to find that both the between-regions and within-region components of the inequality measure were also steadily decreasing. One more interesting finding is that the share of between-region and within-region inequality almost did not change except a small change in the 2004 sample compared with the other two samples: 1998 and 2002.

Table 4.16. Theil's second measure and its decomposition for health status

	1998		2002		2004	
	Index	Share	Index	Share	Index	Share
Between-region	0.079	54 %	0.075	54 %	0.072	53.3 %
Within-region	0.068	46 %	0.065	46 %	0.063	46.7 %
Total inequality	0.147		0.140		0.135	

Looking at the composition of the inequality we find that the health differences were larger between regions than within regions. Figure 4.10 below also shows the steadily downwards trend of Theil's measure and its decomposition. The figure further supports our findings about the steady decline of the Theil's second measure of spatial regional inequality in health status in the 1998-2004 period.

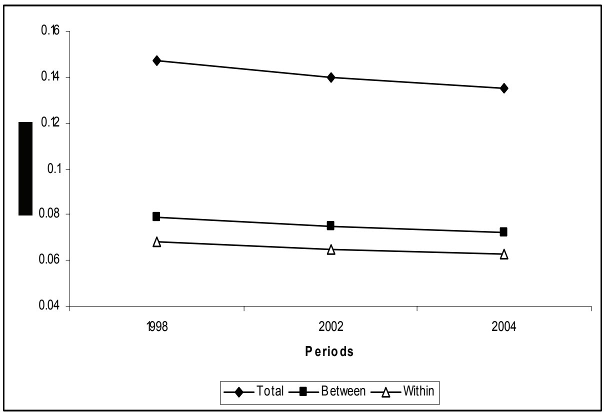


Figure 4.10. Trends of Total, Between- and Within-Region Inequality for Health Status

4.6.2.3. A comparison of spatial patterns between healthcare facilities and health status

Figure 4.11 displays the clusters of favored regions with respect to the composite index of healthcare facilities and health status presented in Table 4.19 and Table 4.20 respectively. Looking at the figure we observe that the favored regions with respect to healthcare facilities do not completely coincide with the favored regions with respect to health status. This means that the regions that were better equipped recently with the healthcare facilities may not yield a better health status.

The comparison of the spatial patterns of healthcare facilities with health status for all three years under study leads to the conclusion that for the year 1998, the most favored regions with respect to healthcare facilities were those located in most parts of the North East (Quang Ninh, Phu Tho, Cao Bang, Yen Bai, Lang Son and Thai Nguyen), the Red River Delta (Hanoi, Hai Phong, Vinh Phuc, Ha Nam, Nam Dinh and a small part in the South East (Ho Chi Minh City) while the most favored regions with respect to health status were those located mainly in the Red River Delta (Hanoi, Hai Phong, Thai Binh, Hai Duong and Vinh Phuc), the North and South Central Coast (Ha Tinh, Thua Thien Hue and Da Nang), the South East (Ho Chi Minh city, Ba Ria-Vung Tau), and some provinces of the Mekong River Delta (Vinh Long and Tien Giang). The similar patterns can also be found in the 2002 and 2004 samples. However, the health status was better in 2002 and 2004 in some parts of the South Central Coast and most parts of the South East and in a large part of the Mekong River Delta.

The differences in the favored regions between healthcare facilities and health status can be explained. The good health status was recorded in most lowland parts of Vietnam which are resided by the Kinh people, and characterized with a higher level of economic development. The most favored regions with respect to health status were in the lowland areas, while the least favored regions with respect to health facilities were recorded in the upland, backward and remote areas resided by many ethnic minority people, and characterized by a lower level of socio-economic development despite the recent increase in the budget share spent on this. These results are in conformity with the practice that the government of Vietnam invested relatively more in the health sector in mountainous and backward areas in order to improve the health status of the people. Despite this effort, the health status in these areas is still lagging behind, probably due to unhealthy lifestyles and living habits that could cause a negative impact on people's health. This finding further consolidates the findings by Wiley and Camacho (1980), and Shirakawa and Morimoto (1991) in their respective studies on the relationship between health and lifestyle in Alameda County (US) and Japan.

+ Healthcare facilities[37]:

The non-contiguity case:

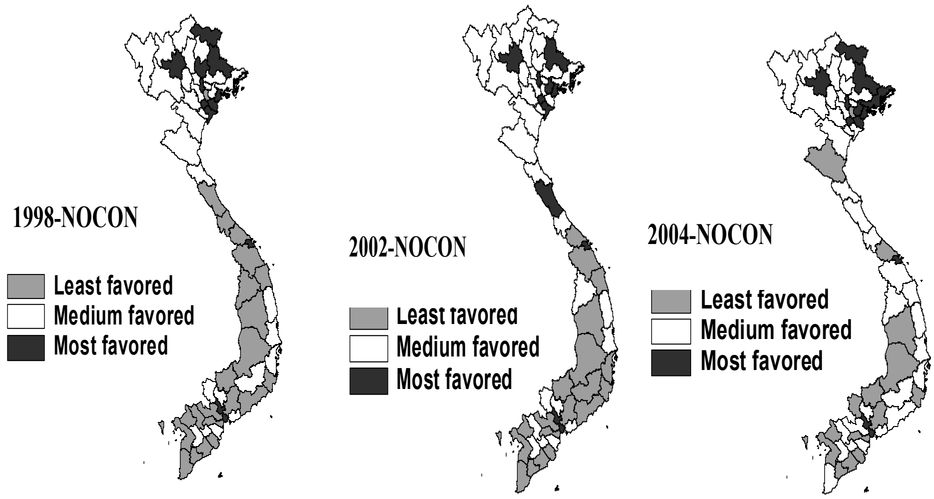
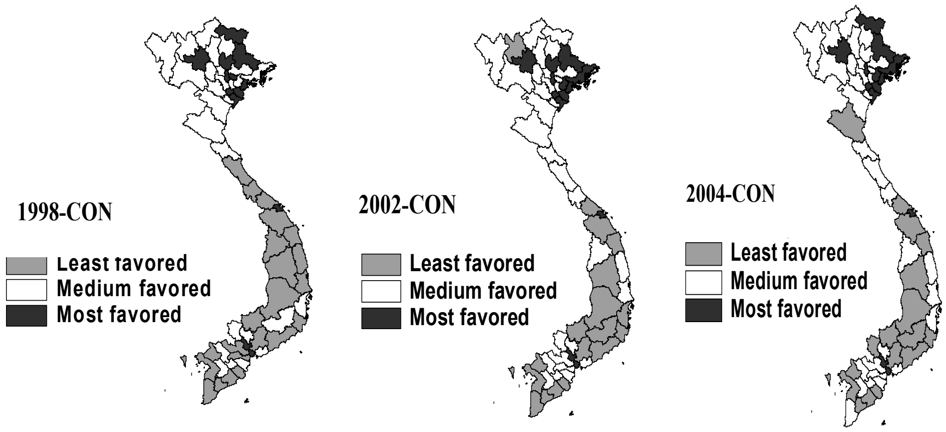


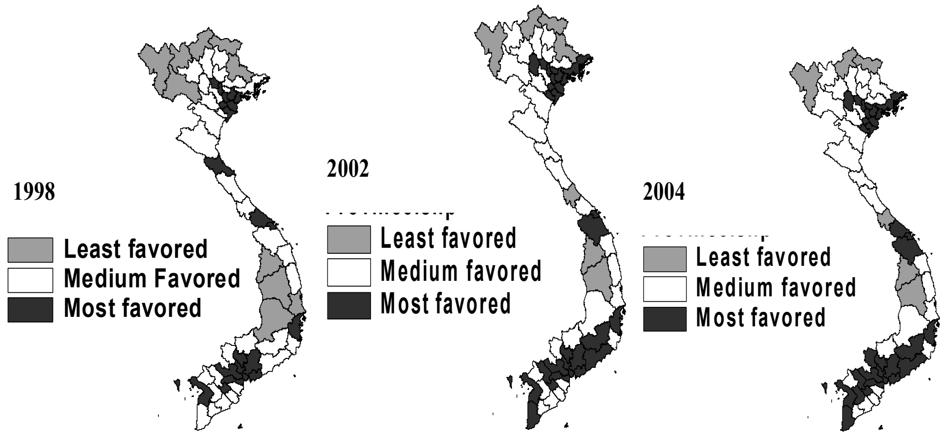
Figure 4.11. Clusters of favored regions with respect to the composite indices of health facilities and health status: a comparison of the spatial patterns

The contiguity case:



[37] The division of favored regions is based on the value of the composite index. For healthcare facilities, most favored regions are those with a value of the composite index larger than 0.55; medium favored regions are those with a value of the composite index between 0.40 and 0.55; and least favored regions are those with a value of the composite index less than 0.40.

+ Health Status[38]**:**



4.7. Testing spatial dependence using spatial econometric model

Spatial dependence models deal with spatial interaction (spatial autocorrelation) and spatial structure (spatial heterogeneity) primarily in cross-section data (Anselin, 1988a, 1999). Spatial dependence models use a metric concept of economic distance (Anselin, 1988 and Conley, 1999). Presently, with the increasing availability of micro and macro level panel data, spatial panel data models studied in Anselin (1988a) become attractive in empirical economic research. To name some, Case (1991); Kelejian and Robinson (1992); Case, Hines and Rosen (1993); Driscoll and Kraag (1998); Baltagi and Li (2001b); and Bell and Bockstael (2000) have applied spatial model for panel data in their researches.

These researches have pointed out that all the data sets those could exhibit spatial correlation as well as heterogeneity across locations or units can be modeled with a spatial error component model (Baltagi and Li, 2001a).

On the background of this short review of the literature and argument, in this section we intend to discuss the spatial dependence between the two dimensions of health studied earlier in this chapter, namely health facilities and health status. In fact, there is a correlation between these two dimensions of health as health facilities might have a positive impact on health status. Thus, an applied model-based approach is taken where spatial econometric methods are introduced to test for spatial correlation between the two dimensions of health. Our goal is to compare the spatial or boundary effects between health facilities (both in case of contiguity and of non-contiguity) and health status. To achieve this goal, we aim at specifying several specifications that show a relationship between the two dimensions of health. Then we go on to test for the spatial dependence between health status and health facilities. Based on the data available, we specify the following regression models for spatial dependence testing:

1. The composite index with respect to health status is a function of the composite index with respect to health facilities (for both the contiguity and non-contiguity cases),

[38] For health status, most favored regions are those with a value of the composite index larger than 0.801; medium favored regions are those with a value of the composite index in the range of 0.761-0.801; and least favored regions are those with a value of the composite index smaller than 0.761.

2. The composite index with respect to health status is a function of childhood immunization, hospital beds, medical doctors and pharmacists. For this specification, we will also test the spatial dependence for both the contiguity and non-contiguity cases of health facilities.

The rationale for the selection of the above two specifications for spatial dependence among the dimensions of health needs to be discussed herewith. As we argued earlier, there was a direct relationship between health facilities and health status of people. Good health facilities would yield good health records for people. In addition, due to contiguity effects, health facilities in one region would have an effect on health status of people living both in that region and the neighboring ones in the first order contiguity assumption made earlier. Therefore, the first specification would help to explore the direct impact of health facilities on health status, and the second specification would consider the impact of each health facilities' indicator on health status in order to help us drawing a panorama of spatial dependence between the two dimensions of health.

Thus, next section will be used to explore the spatial dependence between the two dimensions of health under the two models specified above. We structure this section as follows. The theoretical model and test statistics for spatial dependence is given in section 4.7.1. The empirical model and data for the analysis are discussed in section 4.7.2. Finally, the empirical results are provided in section 4.7.3.

4.7.1. Model and test statistics for spatial dependence

4.7.1.1. The theoretical models

Spatial dependence in a collection of sample data observation refers to the fact that one observation associated with a location which might label *i* depends on other observations at location *j* $(j \neq i)$. Formally, we might write (LeSaga, 1999):

$$y_i = f(y_j), i = 1, \ldots\ldots, n; j \neq i \qquad (4.5)$$

Note that we allow the dependence to be among several observations, as index *i* can take on any value from *1* to *n*. We expect sample data observed at one point in space to be dependent on values observed at other locations for two reasons. Firstly, data collection of observations associated with spatial units such as provinces might reflect measurement error. This would occur if the administrative boundaries for collecting information do not accurately reflect the nature of the underlying process generating the sample data. Secondly, and probably a more important reason is that the spatial dimension of socio-demographic, economic or regional activities may truly be an important aspect of the modeling problem (LeSage, 1999). Regional science is based on the premise that location and distance are important forces at work in human action. These notions have been formalized in regional science theory that relies on notions of spatial interaction and diffusion effects, hierarchies of place and spatial spillovers.

However, spatial correlation as expressed in model (4.5) may not be of much interest (Case, 1991). For example, health status in a province in the north of Vietnam would not have anything to do with the health status in a province located in the central part of the country. This means that the health status among provinces does not have any effect on each other. Instead, it would be more useful to test for spatial correlation between health dimensions such as health status and health facilities. Based on these arguments, the model under the null hypothesis of no spatial dependence in the standard linear regression model is:

$$y = X\beta + u \qquad (4.6)$$

where y is a (N,1) vector of observations on a dependent variables taken at each of N locations, X is a non-stochastic (N, k) matrix of explanatory variables, β is a (K,1) vector of coefficients and u is a (N,1) vector of independent and identically distributed disturbance terms from the OLS regression. N is the number of observations or locations. Under the alternative of spatial error autocorrelation, the u is no longer independent and identically distributed (Anselin 1988). The disturbance vector u in equation (4.6) is assumed to have spatially auto-correlated disturbances. The spatial auto-correlated model in the error term is then defined as:

$$u = \lambda Wu + \mu \qquad (4.7)$$

where λ is a (N,1) vector of the spatial autoregressive coefficient that implies province or border effects, W is a *(N,N)* symmetric spatial matrix and μ is a (N,1) vector of independent and identically distributed errors and is independent of u. The spatial matrix W expresses the strength of potential interaction between each observation. The spatial matrix W is constructed based on simple binary contiguity coefficients as w_{ij} equals one if i and j have a common border or 0 otherwise (see, among others, Anselin 1991, and Baltagi and Li 2001a, b).

In regression analysis, the weights matrix W is typically used in row-standardized form, i.e. the row elements sum to one (Anselin, 1991). As a consequence the matrix W will be symmetric and its eigenvalues will be less than one in absolute value. Obviously, the proper choice of a spatial weights matrix is an important problem in spatial analysis, and one which has not yet obtained a satisfactory solution (Anselin, 1988b). To simplify the analysis, in our analysis we therefore assume the matrix is a non-stochastic matrix.

However, the linear model presented in equation (4.6) and (4.7) is just a data generating process of spatial dependence. Subsequent to this we need to determine whether substantive spatial dependence (i.e., an autoregressive residual pattern due to the omission of a spatial lag) or a nuisance type of spatial dependence (i.e., an autoregressive error structure) occurs. These types of misspecification are likely to show up in misspecification testing if the true data generating processes are either the spatial AR error model, which is given by:

$$y = X\beta + (1 - \lambda W)^{-1} u \qquad (4.8)$$

or the spatial lag model, given by:

$$y = \rho Wy + X\beta + \varepsilon \qquad (4.9)$$

where:

W is spatial weight matrix defined above representing the spatial morphology, and λ and ρ are scalar autoregressive parameters.

The test used to identify spatial dependence is Moran I statistics, and the diagnostic tools used to identify a spatially autoregressive error term or an omitted spatial lag are the

Lagrange Multiplier (*LM)* tests (see, among others, Anselin, 1988 and Burridge, 1980). We will discuss the tests in the next section.

4.7.1.2. The diagnostic testing method

Estimation and testing using panel data models have been extensively studied (see Baltagi 2001). In this study we apply several diagnostic tests for spatial correlation such as Moran *I* test and Lagrange Multiplier (*LM*) test to test for the spatial dependence in the regression's error in order to arrive at the conclusion of whether there is a presence of spatial correlation between the dimensions of health namely health status and health facilities. These tests are very important, because they help us to arrive at a correct specification for the relationship and analysis of the spatial effects between health facilities and health status. In addition, if we ignore spatial correlation due to random regional effects, it would result in inefficient estimates and misleading inference about the relationship between health facilities and health status.

As pointed out earlier, the focus of the test for spatial dependence is on whether the λ in equation (4.7) equals zero based on either the Moran I (*MI*) statistic of the OLS residuals of (4.6) (Moran, 1950; Cliff and Ord, 1972, 1973, 1981; Anselin, 1988a) , or the *LM* maximum likelihood test by a spatial error model or spatial lag model presented in the form of equations (4.8) and (4.9) (see, Florax *et al*., 2002; Ord, 1975). The Moran I statistic can be used with equation (4.7) to measure covariance in errors between joining provinces relative to the variance errors in a given province. The *MI* statistic for the residuals of model (4.7) is defined as follows:

$$MI = (N/S)e'We/e'e \qquad (4.10)$$

where e are the OLS residuals, N is the number of observations and S is the sum of all elements in the spatial weights matrix W. For a row-standardized W, S simplifies to N, so that $N/S=1$ and can be ignored. Therefore (4.10) becomes[39]:

$$MI = e'We/e'e \qquad (4.11)$$

For normal error terms the distribution of the standardized Moran I statistic is shown to be asymptotically distributed as χ^2 with one degree of freedom.

The exact distribution of Moran I statistic (4.11) can be derived for normal errors. The statistic shows a striking similarity to the familiar Durbin-Watson test. In the spatial case, the critical value will depend not only on X, but also on the asymmetric matrix of W, which is clearly not very practical (Anselin, 1991). Also, Cliff and Ord (1972, 1973, 1981) show that the asymptotic distribution of Moran *I* based on least-square residuals corresponds to a standard normal distribution.

When the spatial regression models are estimated by maximum likelihood (ML), inference on the spatial autoregressive coefficients may be based on Wald or asymptotic t-test (from the asymptotic variance matrix) or on a likelihood ratio test (see Anselin 1988, chapter 6; Anselin and Bera 1998). Both approaches require that the alternative model (i.e., the spatial model) be estimated. In contrary, a series of test statistics based on the Lagrange

[39] For details of the proofs, please refer to Anselin (1988) and LeSage (1999).

Multiplier (LM) principle require only estimation of the model under the alternative. LM test also allow for the distinction between a spatial error and a spatial lag alternative.

In addition, In Monte Carlo experiments Anselin and Florax (1995) have also shown that the Bera-Yoon LM tests have high power in discriminating between spatial lag and error dependence. The basic idea of these tests consists in correcting the LM test statistics for spatial error and lag dependence. According to Anselin and Florax (1995: 25-29), LM(err) and LM(lag) can be presented as:

$$LM_{(err)} = (e^{'}.W.e/\hat{\sigma}^2)^2 / tr(W^{'}W + W^2) \quad (4.12a)$$

and

$$LM_{(lag)} = (e^{'}.W.y_T/\hat{\sigma}^2)^2 / [(Wy_0\hat{B})^{'} MWy_0\hat{B}/\hat{\sigma}^2 + tr(W^{'}W + W^2)] \quad (4.12b)$$

with the ML estimator $\hat{\sigma}^2 = e^{'}.e/n$ for the error variance σ^2 and $M = I - y_0(y_0^{'}y_0)^{-1}y_0^{'}$ known from regression analysis as a projection matrix. To distinguish among alternatives, the spatial lag dependence is eliminated from LM(err) by extracting a function of $e^{'}.W.y_T$, whereas a function of $e^{'}.W.e$ is subtracted from LM(lag) in order to control spatial error dependence (see, among others, Bera and Yoon, 1993; Anselin and Florax, 1995). This means that the adjusted LM error test (LM(err) robust test) responds to spatial error dependence but not to spatial lag dependence. In contrary, the adjusted LM lag test (LM(lag) robust test), is expected to indicate spatial lad dependence but not spatial error dependence. Under the assumption that the errors are normally distributed, both the LM(err) and LM(lag) as well as their robust counterparts tests follow a χ^2 distribution with one degree of freedom.

It is to be noted that the three different estimators are appropriate depending on the model specification. First, the estimators using OLS method are appropriate for the specification with an independent and identically distributed error term and without spatially lagged dependent variable. Second, a maximum likelihood estimators are appropriate for the spatially autoregressive error model (*MLERROR*). Finally, maximum likelihood estimators are appropriate for the model including a spatially lagged dependent variable (*MLLAG*). For details on the estimation method procedure, see Anselin (1988), and Anselin and Hudak (1992).

4.7.2. Empirical models and data

We apply the spatial error model to panel data to test for spatial dependence in the residuals of the regression. Earlier, we have analyzed spatial inequality in healthcare facilities and health status for the three samples separately. We therefore take this into account in testing for spatial dependence. We specify two linear regression models based on the data generating model (4.6) as follows:

+ *Model I*: $y_{hs} = x_{hf}\beta + u$ (4.13)

+ *Model II*: $y_{hs} = x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + x_4\beta_4 + u$ (4.14)

where: - y_{hs} is N by 1 vector of composite index with respect to health status,
x_{hf} is a N by 1 matrix of composite index with respect to health facilities,

x_1 is a N by 1 vector of variable childhood immunization,
x_2 is a N by 1 vector of variable hospital beds,
x_3 is a N by 1 vector of variable hospital doctors,
x_4 is a N by 1 vector of variable pharmacists, and
N is number of provinces or observations under study.

Below is the summary of the steps to follow in order to identify an appropriate specification for the model:

1. Estimate the initial model in the form of equation (4.6) by means of OLS.
2. Test the null hypothesis of no spatial dependence using *MI* (Moran I statistic test), no spatial dependence due to spatially autoregressive error or due to an omitted spatial lag, using LM_λ and LM_ρ, respectively.
3. If all the tests for the null hypothesis of no spatial dependence are rejected which means that there is no spatial dependent relationship existence in the residuals of the regression, the initial estimates from step 1 are used as the final specification. Otherwise go to step 4.
4. If all the tests for the null hypothesis of no spatial dependence are not rejected, it means that there is existence of the spatial dependent relationship between residuals of regression. In this case, we estimate the specification pointed to by the more significant of the last two tests. For example if $LM_\lambda > LM_\rho$, then estimate the models in the form of equation (4.8) using *MLERROR*. If $LM_\lambda < LM_\rho$, estimate the models in the form of equation (4.9) using *MLLAG*. Otherwise, go on to step 5.
5. If LM_λ is significant but LM_ρ is not, estimate the models in the form of equation (4.8) using *MLERROR*. If LM_ρ is significant and LM_λ is not, then estimate the models in the form of equation (4.9) using *MLLAG*.

The data used for this test are the rescaled and transform data that have been used for the measurement and analysis of the multidimensional regional inequality in health facilities and health status for 1998, 2002 and 2004. We also add the composite index with respect to health facilities and composite index with respect to health status calculated in the previous sections and presented in Table 4.19 and Table 4.20.

4.7.3. Results

We follow the steps presented in section 4.7.2. In Table 4.17 below we report the results of the Moran *I* test statistic and *LM* statistics for the two empirical specifications presented in (4.11), (4.12a) and (4.12b) above. As the results indicate, the *MI* test leads to the acceptance of the null hypothesis of no spatial dependence among regression residuals in model 4.14 for the 2002 and 2004 samples, but the *LM* tests reject the null hypothesis for all cases. According to the *LM* tests, most tests are strongly significant with very low *p*-value ($p<0.01$) (only a few test are significant at 95% significant level) and relatively large LM_λ and LM_ρ values compared to the corresponding critical values of Chi-square at one degree of freedom, we are able to reject the null hypothesis of no spatial dependence among the regression residuals. We therefore accept the hypothesis of spatial dependence among regression residuals. Thus, we conclude that there is spatial dependence among the regression residuals.

*Table 4.17. Results of the spatial dependence test**

Models	Contiguity			Non-contiguity		
	1998	2002	2004	1998	2002	2004
Model 4.13			*N=61*			
Moran I	6.28	5.49	5.57	6.28	5.51	5.62
LM_{λ}	32.48	24.59	25.89	32.77	25.07	26.69
LM_{ρ}	28.05	23.21	24.17	29.14	23.92	25.12
Model 4.14			*N=61*			
Moran I	4.43	3.44	2.81	4.49	3.46	2.81
LM_{λ}	13.09	7.50	4.73**	13.89	8.05	5.26**
LM_{ρ}	16.66	12.92	10.72	20.39	14.69	13.90

* The tests are distributed as Chi-square, with critical values of 3.84 (p=0.05), 6.63 (p=0.01) and 7.88 (p=0.005); ** indicates significant at p=0.05 confident level. The other values are highly significant (p<0.01).

Now, based on Step 4 we will identify the appropriate specification of the models presented in equation (4.13) and (4.14). For model 4.13, for all three years, the test results presented in Table 5.17 indicate that $LM_{\lambda} > LM_{\rho}$, thus we estimate this model as a spatial error model presented in equation (4.8) using *MLERROR*. In contrast, for Model 4.14, as the test results in Table 5.17 show that $LM_{\lambda} < LM_{\rho}$ for all the samples, therefore we estimate this model as a spatial lag model presented in equation (4.9) using *MLLAG*.

The maximum Likelihood estimates of these models using *MLERROR* for Model 4.13 and *MLLAG* for Model 4.14 are displayed in Table 4.18 below. For estimates of the Model 4.13 using the *MLERROR* method, we observe that most estimated coefficients are highly statistically significant at 5% or 1% confidence levels. The coefficients of λ take positive values and stay within the acceptable range of ($-1.139 < \lambda < 1$), which implies that there exists spatial dependence among regression residuals in all three samples (1998, 2002 and 2004) as tested above. Thus, the estimation of the model as a spatial error model using MLERROR is appropriate. Looking at the coefficients of the independent variable we observe that generally speaking there is a positive relationship between the composite index with respect to health status and the composite index with respect to health facilities.

Table 4.18. Maximum likelihood estimates[@]

Models/explanatory variables	Contiguity case			Non-contiguity case		
	1998	2002	2004	1998	2002	2004
Model 4.13 (MLERROR method)			*N=61*			
Constant	0.749 (0.014)	0.763 (0.014)	0.763 (0.013)	0.751 (0.014)	0.766 (0.014)	0.766 (0.013)
Composite index of health facilities	0.088 (0.024)	0.062[B] (0.026)	0.069 (0.022)	0.081 (0.023)	0.056* (0.026)	0.065 (0.022)
LM Lamda	0.723 (0.092)	0.667 (0.106)	0.675 (0.104)	0.731 (0.091)	0.675 (0.104)	0.685 (0.102)
Model 4.14 (MLLAG method)			*N=61*			
Constant	0.261 (0.089)	0.304 (0.099)	0.298 (0.104)	0.241 (0.093)	0.305 (0.102)	0.279* (0.124)
Childhood immunization	0.070[D] (0.047)	0.032[N] (0.043)	0.074[N] (0.056)	0.051[N] (0.048)	0.007[N] (0.045)	0.046[N] (0.078)
Hospital beds	-0.149 (0.047)	-0.085[C] (0.052)	-0.067[C] (0.041)	-0.149 (0.047)	-0.079[D] (0.053)	-0.068[C] (0.041)
Hospital doctors	0.062* (0.028)	0.026[N] (0.027)	0.009[N] (0.021)	0.058* (0.028)	0.022[N] (0.031)	0.007[N] (0.021)
Pharmacists	0.039[B] (0.016)	0.037* (0.017)	0.049 (0.016)	0.045 (0.014)	0.039[B] (0.016)	0.052 (0.014)
LM Rho	0.579 (0.112)	0.569 (0.126)	0.529 (0.129)	0.621 (0.107)	0.594 (0.121)	0.581 (0.122)

[@] dependent variable is the *composite index with respect to health status*; standard errors are in parentheses; superscripts: [B] indicates coefficients are significant at 1% level; * indicates coefficients are significant at 5% level; C indicates coefficients are significant at 10% level; D indicates coefficients are significant at 15% level; and *N* indicates coefficients are not statistically significant. The remaining coefficients, those without superscript are highly statistically significant.

For estimates of Model 4.14 using the *MLLAG* method, we observe that the coefficients of ρ take positive values that stay within the acceptable range of ($-1{,}139 < \rho < 1$), which implies that the spatial lag model specification for the model 4.14 is correct in all samples of 1998, 2002 and 2004 as tested above. The sign of coefficients of the independent variables indicates that in all three samples, variables such as childhood immunization, hospital doctors and pharmacists have a positive impact on the health status of the people. Among these variables, the coefficient of childhood immunization is not statistically significant in all the sample years except for the case of contiguity in 1998. The coefficient of the hospital doctor variable is also not statistically significant for most cases (see Table 4.18). It is surprising to find that the hospital bed variable has a significantly negative impact on health status of the people because we argued earlier that hospital bed is a positive variable. However, this finding is in line with the practice as more hospital beds mean more patients. The people's health status is largely connected with people's health. Thus, number of hospital beds is inversely proportional with the health status of people. This is why the hospital bed variable has a negative impact on health status.

Table 4.19. Composite indices for healthcare facilities and their ranking order by provinces†

Provinces	1998		2002		2004	
	Non contiguity	Contiguity	Non contiguity	Contiguity	Non contiguity	Contiguity
	Comp. index & ranking	Comp. index & ranking	Comp. index & ranking	Comp. index & ranking	Comp. index & ranking	Comp. index & ranking
Hanoi	0.773 (2)	0.800 (1)	0.693 (3)	0.730 (3)	0.772 (2)	0.803 (2)
Hai Phong	0.669 (4)	0.667 (4)	0.627 (4)	0.637 (4)	0.625 (6)	0.635 (6)
Vinh Phuc	0.498 (17)	0.529 (15)	0.468 (27)	0.506 (21)	0.486 (22)	0.522 (15)
Ha Tay	0.419 (32)	0.466 (23)	0.452 (30)	0.506 (22)	0.419 (39)	0.474 (28)
Bac Ninh	0.528 (12)	0.551 (12)	0.493 (21)	0.525 (16)	0.412 (42)	0.446 (34)
Hai Duong	0.508 (15)	0.537 (13)	0.567 (6)	0.602 (5)	0.550 (11)	0.583 (8)
Hung Yen	0.356 (47)	0.415 (34)	0.400 (42)	0.462 (30)	0.381 (50)	0.439 (37)
Ha Nam	0.705 (3)	0.729 (3)	0.706 (2)	0.747 (2)	0.644 (4)	0.683 (4)
Nam Dinh	0.571 (10)	0.607 (8)	0.547 (10)	0.590 (6)	0.541 (12)	0.579 (9)
Thai Binh	0.580 (8)	0.613 (7)	0.519 (16)	0.563 (8)	0.554 (9)	0.595 (7)
Ninh Binh	0.526 (14)	0.535 (14)	0.502 (20)	0.522 (17)	0.476 (27)	0.492 (25)
Ha Giang	0.414 (35)	0.404 (37)	0.452 (29)	0.449 (33)	0.477 (26)	0.472 (29)
Cao Bang	0.577 (9)	0.567 (10)	0.538 (12)	0.538 (15)	0.562 (8)	0.562 (11)
Lao Cai	0.448 (25)	0.437 (29)	0.402 (41)	0.400 (44)	0.474 (29)	0.467 (31)
Bac Can	0.526 (13)	0.528 (16)	0.503 (19)	0.516 (19)	0.498 (18)	0.511 (18)
Lang Son	0.622 (5)	0.616 (6)	0.553 (9)	0.558 (10)	0.636 (5)	0.641 (5)
Tuyen Quang	0.432 (29)	0.450 (25)	0.453 (28)	0.476 (27)	0.450 (31)	0.472 (30)
Yen Bai	0.563 (11)	0.557 (11)	0.559 (8)	0.562 (9)	0.570 (7)	0.571 (10)
Thai Nguyen	0.597 (7)	0.616 (5)	0.526 (14)	0.554 (11)	0.494 (19)	0.524 (14)
Phu Tho	0.406 (38)	0.421 (32)	0.431 (35)	0.454 (32)	0.412 (41)	0.432 (39)
Bac Giang	0.462 (22)	0.492 (17)	0.411 (40)	0.448 (34)	0.475 (28)	0.513 (17)
Quang Ninh	0.482 (19)	0.484 (19)	0.544 (11)	0.551 (12)	0.553 (10)	0.560 (12)
Lai Chau	0.430 (31)	0.412 (35)	0.506 (17)	0.493 (23)	0.446 (33)	0.430 (41)
Son La	0.500 (16)	0.485 (18)	0.529 (13)	0.522 (18)	0.504 (17)	0.495 (22)
Hoa Binh	0.447 (26)	0.452 (24)	0.506 (18)	0.516 (20)	0.508 (15)	0.514 (16)
Thanh Hoa	0.495 (18)	0.481 (21)	0.492 (22)	0.487 (24)	0.470 (30)	0.462 (33)
Nghe An	0.436 (27)	0.420 (33)	0.449 (31)	0.440 (37)	0.394 (47)	0.385 (50)
Ha Tinh	0.417 (33)	0.401 (39)	0.425 (37)	0.418 (41)	0.444 (35)	0.434 (38)
Quang Binh	0.342 (50)	0.326 (53)	0.560 (7)	0.548 (13)	0.487 (20)	0.474 (27)
Quang Tri	0.388 (42)	0.377 (43)	0.432 (34)	0.427 (40)	0.439 (36)	0.432 (40)
T.Thien-Hue	0.317 (54)	0.313 (54)	0.327 (54)	0.327 (55)	0.365 (52)	0.364 (53)
Da Nang	0.608 (6)	0.593 (9)	0.588 (5)	0.580 (7)	0.726 (3)	0.714 (3)
Quang Nam	0.341 (51)	0.338 (50)	0.363 (33)	0.366 (49)	0.397 (46)	0.397 (46)
Quang Ngai	0.362 (46)	0.352 (47)	0.380 (47)	0.376 (47)	0.402 (44)	0.396 (48)
Binh Dinh	0.406 (37)	0.390 (42)	0.446 (32)	0.437 (38)	0.505 (16)	0.492 (24)
Phu Yen	0.400 (39)	0.391 (41)	0.486 (23)	0.480 (26)	0.484 (24)	0.477 (26)
Khanh Hoa	0.452 (23)	0.440 (28)	0.390 (44)	0.388 (45)	0.444 (34)	0.440 (36)
Kon Tum	0.354 (48)	0.345 (49)	0.477 (25)	0.468 (29)	0.434 (37)	0.425 (42)
Gia Lai	0.311 (56)	0.305 (56)	0.327 (55)	0.327 (56)	0.328 (55)	0.329 (55)
Dak Lak	0.276 (57)	0.273 (57)	0.245 (60)	0.245 (61)	0.261 (61)	0.261 (61)
Lam Dong	0.407 (36)	0.401 (38)	0.383 (46)	0.383 (46)	0.400 (45)	0.399 (45)
Ho Chi Minh	0.777 (1)	0.789 (2)	0.773 (1)	0.798 (1)	0.796 (1)	0.819 (1)
Ninh Thuan	0.335 (52)	0.328 (52)	0.362 (49)	0.359 (50)	0.388 (49)	0.383 (51)
Binh Phuoc	0.248 (60)	0.246 (61)	0.242 (61)	0.246 (60)	0.287 (60)	0.287 (60)
Tay Ninh	0.449 (24)	0.447 (26)	0.469 (26)	0.475 (28)	0.353 (54)	0.361 (54)
Binh Duong	0.472 (21)	0.484 (20)	0.519 (15)	0.540 (14)	0.523 (13)	0.539 (13)
Dong Nai	0.366 (44)	0.372 (44)	0.355 (51)	0.368 (48)	0.379 (51)	0.391 (49)
Binh Thuan	0.366 (45)	0.355 (46)	0.359 (50)	0.355 (51)	0.402 (43)	0.396 (47)
B.Ria-V. Tau	0.482 (20)	0.472 (22)	0.485 (24)	0.484 (25)	0.512 (14)	0.507 (20)
Long An	0.390 (41)	0.397 (40)	0.396 (43)	0.411 (42)	0.433 (38)	0.445 (35)
Dong Thap	0.396 (40)	0.407 (36)	0.446 (33)	0.461 (31)	0.478 (25)	0.493 (23)

An Giang	0.329 (53)	0.333 (51)	0.327 (56)	0.335 (53)	0.321 (56)	0.322 (57)
Tien Giang	0.345 (49)	0.361 (45)	0.387 (45)	0.408 (43)	0.388 (48)	0.411 (43)
Vinh Long	0.433 (28)	0.445 (27)	0.418 (39)	0.441 (36)	0.487 (21)	0.508 (19)
Ben Tre	0.432 (30)	0.435 (30)	0.418 (38)	0.431 (39)	0.449 (32)	0.464 (32)
Kien Giang	0.315 (55)	0.311 (55)	0.274 (59)	0.275 (59)	0.297 (59)	0.302 (59)
Can Tho	0.417 (34)	0.422 (31)	0.428 (36)	0.442 (35)	0.485 (23)	0.500 (21)
Tra Vinh	0.244 (61)	0.257 (60)	0.280 (57)	0.295 (57)	0.307 (57)	0.323 (56)
Soc Trang	0.265 (58)	0.268 (58)	0.277 (58)	0.288 (58)	0.304 (58)	0.315 (58)
Bac Lieu	0.261 (59)	0.262 (59)	0.328 (53	0.332 (54)	0.361 (53)	0.365 (52)
Ca Mau	0.368 (43)	0.351 (48)	0.350 (52)	0.341 (52)	0.415 (40)	0.404 (44)

† Rankings are given in parentheses

Table 4.20. Composite indices for health status and their ranking order by provinces†

Provinces	1998	2002	2004
Hanoi	0.835 (2)	0.837 (1)	0.839 (1)
Hai Phong	0.818 (9)	0.821 (7)	0.824 (7)
Vinh Phuc	0.809 (12)	0.812 (11)	0.815 (11)
Ha Tay	0.790 (35)	0.796 (35)	0.800 (36)
Bac Ninh	0.801 (21)	0.807 (19)	0.810 (20)
Hai Duong	0.818 (8)	0.823 (6)	0.826 (6)
Hung Yen	0.807 (14)	0.811 (14)	0.814 (14)
Ha Nam	0.821 (7)	0.825 (5)	0.828 (5)
Nam Dinh	0.828 (5)	0.809 (17)	0.812 (17)
Thai Binh	0.836 (1)	0.811 (13)	0.814 (15)
Ninh Binh	0.801 (22)	0.806 (20)	0.810 (21)
Ha Giang	0.703 (60)	0.715 (60)	0.722(60)
Cao Bang	0.731 (58)	0.742 (58)	0.749 (58)
Lao Cai	0.760 (54)	0.770 (53)	0.775 (52)
Bac Can	0.765 (51)	0.774 (50)	0.779 (50)
Lang Son	0.748 (56)	0.757 (55)	0.763 (55)
Tuyen Quang	0.768 (50)	0.775 (49)	0.779 (49)
Yen Bai	0.770 (49)	0.778 (47)	0.782 (48)
Thai Nguyen	0.788 (41)	0.794 (41)	0.798 (40)
Phu Tho	0.799 (25)	0.804 (23)	0.807 (24)
Bac Giang	0.771 (47)	0.778 (46)	0.783 (46)
Quang Ninh	0.800 (24)	0.805 (22)	0.808 (23)
Lai Chau	0.740 (57)	0.751 (57)	0.751 (57)
Son La	0.761 (53)	0.770 (51)	0.776 (51)
Hoa Binh	0.770 (48)	0.778 (48)	0.782 (47)
Thanh Hoa	0.785 (43)	0.792 (43)	0.795 (43)
Nghe An	0.795 (31)	0.797 (32)	0.801 (31)
Ha Tinh	0.804 (18)	0.796 (34)	0.800 (34)
Quang Binh	0.775 (45)	0.782 (45)	0.786 (45)
Quang Tri	0.774 (46)	0.755 (56)	0.760 (56)
T.Thien- Hue	0.803 (20)	0.799 (30)	0.802 (30)
Da Nang	0.832 (4)	0.836 (3)	0.838 (3)
Quang Nam	0.798 (26)	0.803 (24)	0.807 (25)
Quang Ngai	0.788 (40)	0.794 (40)	0.798 (38)
Binh Dinh	0.790 (37)	0.796 (36)	0.800 (35)
Phu Yen	0.761 (52)	0.770 (52)	0.775 (53)
Khanh Hoa	0.806 (17)	0.811 (15)	0.814 (13)
Kon Tum	0.693 (61)	0.704 (61)	0.712 (61)
Gia Lai	0.730 (59)	0.740 (59)	0.748 (59)
Dak Lak	0.755 (55)	0.764 (54)	0.772 (54)
Lam Dong	0.797 (29)	0.802 (26)	0.806 (27)
Ho Chi Minh	0.834 (3)	0.837 (2)	0.838 (2)
Ninh Thuan	0.785 (44)	0.791 (44)	0.795 (44)
Binh Phuoc	0.789 (39)	0.795 (39)	0.798 (41)

Tay Ninh	0.794 (32)	0.799 (29)	0.802 (29)
Binh Duong	0.806 (16)	0.810 (16)	0.813 (16)
Dong Nai	0.803 (19)	0.808(18)	0.811 (18)
Binh Thuan	0.797 (28)	0.802 (27)	0.805 (28)
B.Ria-V. Tau	0.824 (6)	0.828 (4)	0.830 (4)
Long An	0.809 (13)	0.812 (10)	0.815 (10)
Dong Thap	0.800 (23)	0.806 (21)	0.809 (22)
An Giang	0.790 (36)	0.796 (37)	0.799 (37)
Tien Giang	0.811 (11)	0.815 (9)	0.817 (9)
Vinh Long	0.817 (10)	0.821 (8)	0.823 (8)
Ben Tre	0.792 (33)	0.797 (31)	0.800 (32)
Kien Giang	0.806 (15)	0.812 (12)	0.815 (12)
Can Tho	0.796 (30)	0.801 (28)	0.810 (19)
Tra Vinh	0.791 (34)	0.797 (33)	0.800 (33)
Soc Trang	0.787 (42)	0.793 (42)	0797 (42)
Bac Lieu	0.789 (38)	0.795 (38)	0.978 (39)
Ca Mau	0.798 (27)	0.803(25)	0.806 (26)

† Rankings are given in parentheses

4.8. Conclusion and discussion

Chapter four has dealt with the evolution of regional inequality in healthcare facilities and health status in Vietnam during 1998-2002 and 2002-2004. We applied the multidimensional approach to measure and analyze regional inequality using Theil's second measure. As we considered multi indicators on health in the analysis, the principal component analysis has been applied. We applied the partial common principal component model (PCPC) to estimate common weights for the indicators with respect to healthcare facilities and health status separately. Using the PCPC model we assumed that the first common component is shared by two or more sample periods. We have also separately analyzed the regional inequality for the health facilities under two cases: contiguity and non-contiguity in its first order contiguity in order to compare the regional inequality between the two cases for examining the impact of spillover effects.

The results showed that the total regional inequality has different patterns for healthcare facilities and health status. Inequality in both healthcare facilities and health status moderately decreased in Vietnam from 1998 to 2004.

The results also proved the importance of spillover effects as the healthcare services seemed to be more equally distributed in the case of contiguity than it did in the case of non-contiguity as Theil's second measure in the former was lower than in the latter. Thus, we conclude that if we ignore the existence of spillover effects, the regional inequality will be exaggerated. Therefore, the findings suggest that it is important to take the contiguity effects into consideration.

These were interesting findings and contrast with our initial speculation that the *doi moi* in Vietnam may have deteriorated healthcare facilities and health status between regions in the country because under socialism in the past, people were able to equally get free access to health services. However the *doi moi*, which have increased the national income and then the national income per capita, have made both the country and people wealthier. Therefore the government could spend more to improve infrastructure of the health sector, and people were more willing to spend more on keeping good health by improving their daily diet, having more regular health checkups and spending more on medication in case of contracting disease. Statistics from the household expenditure surveys showed that household expenditure on health has increased by almost 30% during the 1998-2004 period (GSO, 1998, 2002, 2004).

The chapter also reported the tests for spatial dependence among health components. The tests have been conducted for several specifications of the relationship between the composite index with respect to health status and the composite index with respect to health facilities, and between the composite index with respect to health status and the indicators of health facilities. The results of the test confirmed the existence of the spatial effects between the two dimensions of health. The test proved that the spatial error model is an appropriate specification for the former specification while the spatial lagged model is a more appropriate specification for the latter.

Generally, the above findings are in conformity with the health status and health facilities in practice as well as the two dimensions of health in Vietnam have also been spatially correlated as the findings proved. In addition, the research also draws a conclusion about the relation between healthcare facilities and health status as the findings reveal that the regions with a high composite index with respect to health status do not completely coincide with those having a high composite index with respect to healthcare facilities. This means that regions that were better equipped with health facilities (in other words, the regions that get more investment or a higher share of budget spent on health) may not have yielded a better status of general health of people. These regions are the ones with a backward infrastructure and poor socio-economic level of development, and are predominantly mountainous and backward regions. Furthermore, with the short time span of six years, the budgets spent on health may not have enough time to reap the positive change in people's health status. In addition, in the mountainous and backward regions, lifestyle and living habits of people (such as smoking, alcohol consumption, having a breakfast, hours of sleep, hours of work, physical exercise, nutritional balance, and mental stress) were considered as unhealthy. While most lowland regions with a lower share of the budget spent on health already had good records of health status. All these factors contribute to the non-existence of a correlation between health facilities and health status among regions in Vietnam.

Chapter 5
Education Facilities and Enrolments

5.1. Introduction

Education is recognized as an important element of the quality of life. Like health, education is an important dimension of wellbeing. Education has numerous consequences for individuals and society. Usually, people want to learn and acquire knowledge. Those who acquire additional schooling generally earn more over their lifetime, achieve higher levels of employment, get more confidence and political equity, and have more satisfactory careers. Thus, education has significant effects on people's opportunities for a better life (The World Bank, 2006). Education also helps people getting future progress and better achievements. In the developing world, education is considered as an important device to help people escape from poverty. At the World Economic Forum in January 2005 in Davos, education was capturing heightened attention. Also at the Forum, the participating leaders ranked education as a leading global concern, recognizing it as a key to beating poverty and improving people's welfare[40].

Recognizing its substantial social benefits, most countries pay great attention to improving education by investing more in education as they know that investment in education has long been seen as a key factor for development, economic growth, poverty reduction, health improvement and information revolution (Riddle, 2004). A nation will be wealthier and stronger if it has a high percentage of literacy and a good system of education. The role of schooling in promoting stability, wealth and society's democracy has been expressed by Friedman (1955):

> "A stable and democratic society is impossible without widespread acceptance of some common set of values and without a minimum degree of literacy and knowledge on the part of most citizens. Education contributes to both. In consequence, the gain from the education of a child accrues not only to the child or to his parents but to other members of the society; the education of my child contributes to other people's welfare by promoting stable and democratic society"

However, there was great concern about the accessibility to and availabilities of education in most countries, and their effects on literacy and enrolments. People in most countries in the developing world are facing unequal opportunities in access to education among regions within a country or between countries that results in regional differences in education's enrolments and literacy rates. Studies by the World Bank and UNDP revealed that the differences in enrolment rates and literacy rates at different levels of education and education facilities such as number of teachers, number of schools among regions have been increasing in most developing countries in Africa, Asia and South America (Table 5.1) that resulted in differences in development levels among regions (World Bank 2006). A study on Vietnam by the World Bank and the General Statistics Office of Vietnam (2005) finds that the regions with a higher budget share on education (mostly poor and backward regions) do not coincide with regions with higher rates of enrolments. This is because investment aiming at improving education largely focused on areas resided mostly by poor and ethnic minority people those are very backward and far behind development

[40] http://findarticles.com/p/articles/mi_m1309/is_1_42/ai_n14693233/print

compared with the Kinh[41] living in lowland areas. Even good education facilities would not be able to draw children in these areas to school. In addition, the drop-out rate is high in these areas.

In the case of Vietnam, during the *doi moi* the Government has set the Millennium Development Goals in which education is considered as a great equalizer of opportunities between rich and poor, between men and women and among regions. But the equalizing promise of education can be realized only if children from different backgrounds and different regions have equal opportunities to benefit from quality education (Vietnam Government, 2005). However, the children’s ability to benefit from school is strongly influenced by the cognitive and social skills they acquire in their early childhood[42] (World Bank, 2006). Thus, the government of Vietnam has put many efforts in order to have an equal distribution of investments in education among regions in order to bring good education to all regions in the country to encourage people getting education in order to achieve a high rate of enrolments for all regions (or for every region) (UNDP and the World Bank, 2006). This chapter will investigate and analyze spatial inequality in education facilities (investments in education) and education enrolments among regions in Vietnam. In addition, the chapter also tries to draw a picture of the relationship between investment in education and its outcome (enrolments) among regions in the country by a visual inspection of the composite indices of both the education facilities and education enrolments.

The remainder of chapter five is as follows. A brief overview of the national education system is provided in section 5.2. The spatial patterns of education are given in section 5.3. Description of the data and choice of indicators are in section 5.4, and the analysis of empirical results is given in section 5.5. Finally, the summary of the chapter is structured in section 5.6.

5.2. A brief overview of the national education system in Vietnam

Education in Vietnam was heavily influenced by Chinese Confucianism until around 1954 when the country was divided. Under Confucianism, education was essential for admission to the ruling class of scholars-officials, the mandarins. At the time from 1954 to 1975, South Vietnam used the United State’s system, while the North had mass education and training following the socialist style with the basis of theories by Karl Marx and Vladimir Lenin. The purpose of the Northern system was to train people to live in a communist society.

Before the 1950s, poverty was a major impediment to learning, and secondary and higher education were beyond the reach of all but a small number of upper-class people. Subsequently, however, the rival regimes in Hanoi and Saigon (now Ho Chi Minh City) broadened the educational opportunities. Both governments accomplished this despite the shortage of teachers, textbooks, education equipment, and classrooms and despite the

[41] Kinh is the Vietnamese ethnic people usually called Viet people in English, who are dominant in Vietnam’s population (making up more than 90% of population of Vietnam).

[42] Evidence suggests that the gains from early interventions (assisting children from low-income families with the Head Start program by providing them comprehensive education, health, nutrition and parent involvement services as done in the US) can dissipate if disadvantaged children go on to low quality primary schools. Currie and Thomas (1995) find that gains in vocabulary and reading test scores faded out among black Head Start children while they were still in elementary grades but not among whites even though initial gains in test scores were the same for both groups.

disruption by the Vietnam War in the 1960s and early 1970s. The school system was originally patterned after the French model, but the curriculum was revised to give more emphasis to Vietnam's history, language and literature and - in Hanoi - to the teaching of revolutionary ethics and Marxism-Leninism.

Table 5.1. Net enrolment ratio in primary education and literacy rate in selected countries (%)[43]

Countries	Indicators	1991	1999	2002	2004
Brazil	Net Enrolment Rate	85.6	92.4	93.0	96.4
	Literacy Rate	*n.a*	94.2	96.8	n.a.
Bangladesh	Net Enrolment Rate	80.1	93.0	94.5	97.6
	Literacy Rate	44.7	*n.a.*	63.6	*n.a.*
India	Net Enrolment Rate	*n.a.*	87.4	*n.a.*	96.1
	Literacy Rate	61.9	*n.a.*	76.4	*n.a.*
Indonesia	Net Enrolment Rate	97.3	97.9	99.6	99.0
	Literacy Rate	96.2	*n.a.*	*n.a.*	98.7
Chile	Net Enrolment Rate	*n.a.*	92.8	*n.a.*	94.1
	Literacy Rate	98.4	n.a.	99.0	*n.a.*
Kenya	Net Enrolment Rate	*n.a.*	64.3	64.2	77.0
	Literacy Rate	*n.a.*	80.3	*n.a.*	83.4
Vietnam	Net Enrolment Rate	84.9	94.4	93.9	94.6
	Literacy Rate	86.1	91.4	91.7	92.2

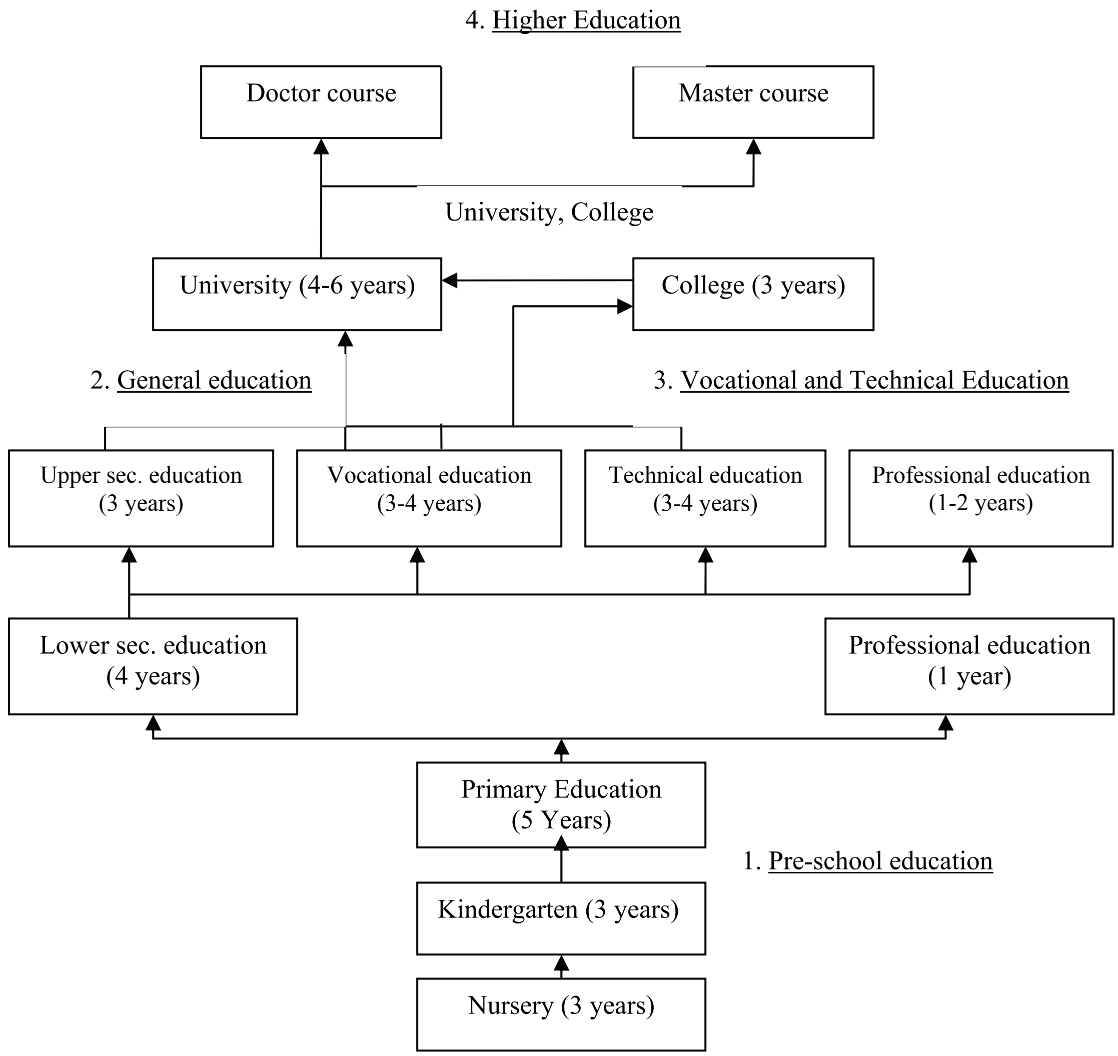
Source: United Nations, Millennium Development Goals Indicators
Notes: *n.a* means not available

After the War ended in 1975, Vietnam was united. The communist system spread throughout the country. Since then, education and training in technology were considered to be equally important as that of teaching communist ideals. After the unification, all public and private schools in the South were taken over by the state as a first step toward integration into a unified socialist school system. Under this united government, the system had been divided into five categories: pre-primary, primary, intermediate, secondary, and higher education. In the new system (effect in 1993), these categories have been sustained but some changes have been made to the divisions of the system. According to the government decision No. 90/CP, dated 24 November 1993, the new system consists of five-sub-systems: pre-school education (pre-primary), general education (including primary, lower and upper secondary), vocational-technical education, higher education (including university education and post graduate education).

Pre-school education is composed of three years for kindergarten starting from three years of age. From six years of age children are enrolled into primary education, which lasts for five years, leading to the certificate of primary education. Upon the completion of primary education, most of the students continue to the basic (lower) secondary education for four years and some of them may be admitted to vocational training for those who could not follow up to secondary education. Finishing basic secondary education with the certificate, some of them can continue specialized secondary

[43] Net enrolment rate in primary education is the number of children of official primary school age who are enrolled in primary education as a percentage of the total children of the official school age population; and literacy rate, as the percentage of the literate population aged 15 years and over, is defined as those with the ability to read and write

education or upper secondary school general for three years or technical for three to four years or vocational training for three to four years based on their education record in lower secondary education. Finishing upper secondary education usually at the age of 18 or after 12 years of schooling, students have to take the national school leaving examination. For those who fail to pass this examination it would be possible to take another exam organized around one month later. If they still fail in the second try they can repeat grade 12 for another year and take the exam again next year. If they don't want to repeat grade 12, they would not get a diploma of general education. With a diploma from upper secondary school or a diploma of general education, a student can take part in the entrance examination for higher education (in university or college). For the undergraduate level, there is either short-term higher education (of three years) or long-term higher education (of four to six years) of regular full-time education or part-time education (service education of continuing education). For post-graduate level, there are Master's programs (of two year duration) and PhD programs (of two to four years length). Figure 5.1 below presents the layout of Vietnam's education system.



Source: Ministry of Education and Training of Vietnam (2006

Figure 5.1. The structure of Vietnam's education system

5.3. Spatial patterns in education facilities, enrolments and literacy rate

Education in Vietnam has been at the centre of attention since the country seized her independence in 1945. Since then many successive governments had considered education a measure to help people getting rid of illiteracy and poverty, and improving their wellbeing. Under the Ho Chi Minh regime the government always aimed at making upland areas to match up with lowlands in development level by focusing on improving education in remote and backward upland areas. Despite all the efforts there were still gaps existing in enrolments, literacy rates, and education facilities among regions in the country.

Before the *doi moi*, education in Vietnam was totally free for all, irrespective of race, religion and ethnic group, etc. Since the *doi moi* inaugurated in 1986, free education was no longer available. Thus, this transition made education no longer free in the country. Instead, children's schoolings were partly paid by their parents. However the government still set to provide education to all by paying more attention to increasing budgets for education in order to equalize basic education to all, irrespective of regions, ethnic groups and social status. In recent years, primary education was again free to all in order to help boosting up education among poor, underprivileged, and in backward and mountainous regions.According to the official data from the General Statistical Office of Vietnam (GSO) there were different trends in education facilities and enrolments. Table 5.2 presents the pupils per teacher ratios and the pupils per classroom ratios in eight regions in the country. The table shows that the ratios steadily declined over time. We also observe that the ratios were higher in lowland areas than in the uplands. The reasons were that in practice population density is higher in the lowlands, which would make the number of pupils higher while people were scattered in the up-lands that would make the number of pupils fewer. In addition, in its policy aiming at providing education to all, the government of Vietnam has paid much attention to building more schools in and assigning more school teachers to upland and backward areas in order to boost up education in these regions. We also observe from Table 5.1 that the decline in the pupils per teacher ratios over time was faster in most lowland and coastal areas than in the up-land ones as in the Red River Delta (lowland) the number of pupils per one teacher went down to 22 in 2002-2004 from 38 in 1990-1991 (decreased by 42%) while in the North West (upland) the decline was only 35% (from 25 to 16).

Table 5.2. Average pupils per teacher ratios and average pupils per classroom ratios in primary and secondary education in eight regions

	1990-91		1994-95		1997-98		1998-99		2001-02		2003-04	
	(1)	*(2)*	*(1)*	*(2)*	*(1)*	*(2)*	*(1)*	*(2)*	*(1)*	*(2)*	*(1)*	*(2)*
Red River Delta	38	41	37	39	33	39	31	39	24	38	22	37
North East	31	33	30	32	28	31	27	31	20	29	17	28
North West	25	27	26	26	25	27	23	26	18	26	16	24
North Central	33	36	35	36	33	36	31	35	25	35	21	34
South Central	35	35	34	35	32	36	32	36	27	36	24	34
Central Highlands	30	32	33	31	33	32	32	33	27	33	25	33
South East	38	37	38	36	35	38	33	37	28	37	25	36
Mekong R. Delta	37	35	37	35	35	35	32	35	26	33	22	32
National average	*35*	*35*	*34*	*35*	*32*	*35*	*30*	*33*	*25*	*34*	*22*	*33*

Sources: Ministry of Education and Training (MOET) and General Statistical Office of Vietnam (GSO).

Notes: (1) indicates number of pupils per one teacher, (2) Indicates number of pupils per one classroom.

Figure 5.2 below reveals that the number of pupils in the lower secondary education and the upper secondary education has increased over time, while the number of pupils at the primary education lightly declined over the years. As the statistics show, the number of pupils at primary education level has declined during the period of the 1995-1996 school year and the 2004-2005 school year (from 10.2 million to 7.7 million). The decline was steady during the period with a light rate year by year. This decline would be caused partly by the large drop-out rates at the primary education (over 7% at national average), and decline in repetition rates at the primary education over time. The decline could also be caused by: *(i)* the decline in the number of children at the age of primary education (age from 6 to 11) as Table 5.3 below reveals that the net enrollment rate increased between 1995 and 2004 (from 91.4% to 94.6%); *(ii)* children left school because their parent could not afford to pay for their education and they need more labor, so they ask their children to leave school in order to work on the field or to earn more income for the household (World Bank Vietnam, various years).

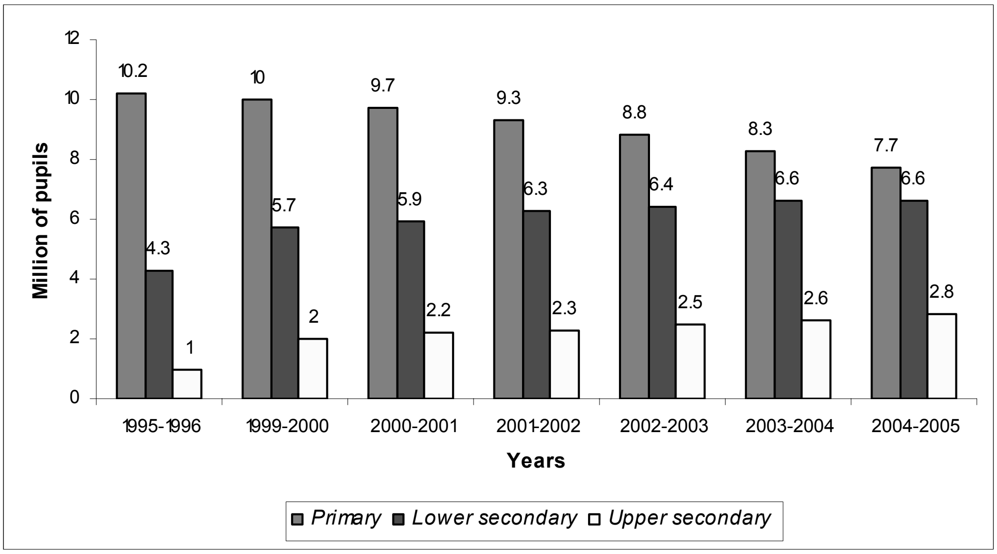


Figure 5.2. Grade of pupils by selected years

Table 5.3. Net enrolment ratio in the age group of 6 to 11 in percent of children of the age group

	1990-91	1994-95	1997-98	1998-99	2001.02	2003-04
Red River Delta	96.4	94.7	98.2	98.0	98.3	98.6
North East	90.6	91.1	96.4	93.9	92.1	93.5
North West	61.3	90.4	92.1	94.8	91.5	92.4
North Central	84.8	92.1	98.1	94.8	94.7	95.2
South Central	95.7	92.3	99.8	97.8	98.2	98.5
Central Highlands	92.5	90.7	91.1	89.2	88.6	88.9
South East	76.4	92.2	97.9	96.1	96.7	97.6
Mekong R. Delta	81.2	87.5	94.7	90.8	91.5	91.9
National average	*84.9*	*91.4*	*96.0*	*94.4*	*93.9*	*94.6*

Sources: MOET and GSO.

The net enrolment rates in eight regions presented in Table 5.3 show that generally speaking the net enrolment rate increased over time, from 84.9% in the 1990-91 school year to 94.6% in the 2003-04 school year. During that period, however, there were ups and downs. For example, in the 1997-1998 school year, the rate has increased to 96% but then decreased to just over 94% in the next school year. This sudden decline can be attributed to changes in the household economy caused by the economic crisis in 1997-1998. Households were getting poorer, that could force parents to make their children leave school. The enrolment rate showed its spatial patterns as the rate was different among regions (see Table 5.3). The highest rate was found mostly in lowland and coastal areas such as in the Red River Delta, the North and the South Central Coast and the South East where big and more developed cities as Hanoi, Hai Phong, Vinh, Hue, Da Nang, Ho Chi Minh City, Vung Tau and Binh Duong, etc. were located. The lower rate of enrolment was found mostly in the mountainous and upland areas in the North, the Central Highlands and most parts of the Mekong River Delta. These regions are characterized by a large portion of population as ethnic minority, poor infrastructure, and remote and far away from most national urban centers.

Table 5.4. Adult literacy rate (in the age group of 15-35) in eight regions (%)

	1990-91	1994-95	1997-98	2001-02	2003-04
Red River Delta	95.07	95.52	95.6	95.8	96.2
North East	68.28	76.87	82.57	89.9	90.0
North West	52.06	55.63	73.55	75.6	76.0
North Central	81.34	86.89	92.14	93.4	93.9
South Central	84.61	88.00	92.06	92.2	92.8
Central Highlands	53.20	67.23	87.1	87.6	88.2
South East	91.15	91.95	92.57	93.2	93.6
Mekong R. Delta	86.08	88.75	89.21	89.4	89.8
National average	*86.1*	*89.9*	*91.4*	*91.7*	*92.2*

Sources: MOET, GSO and Vietnam Human Development Report 2006.

Table 5.4 presents the adult literacy rate in eight regions. According to the data shown in the table, the percentage of the literate population aged 15 years and over, defined as those with the ability to read and write, increased over time (from 86% in 1990-1991 to 92.2% in 2003-2004). We take the age of 15 and over because at this age many people enroll in secondary education, or enter the workforce. Thus, this can be considered a crucial point beyond which the ability to read and write is decisive in many ways. Though the literacy rate in Vietnam is relatively high as the table above shows, sharp regional differences can be seen when looking at spatial distribution of literacy in the country. As in the case of other welfare indicators, such as health facility and health status, the most obvious contrast is between highland and lowland areas. Almost all coastal areas, all of the Red River Delta, and most communes in the South East have a literacy rate above 85%. Also in the Mekong River Delta, the majority of the communes have a literacy rate of at least 85%. Some clusters of communes in the upland areas in the North West and the Central Highlands have rather low rates of literacy, merely above 50% in early 1990s. This is in contrast with the lowland areas where the literacy rate was recorded as high as above

80 % (found in the Mekong River Delta) or above 90% (found in the Red River Delta and in the lowland areas in the North Central Coast and the South Central Coast). On the other hand, a significant area of the North East has a literacy rate above 85%. The highest literacy rates in the country can be found mainly in many areas of the Red River Delta (more than 90%), and particularly also in the central part of the North Central Coast, in Ha Tinh province and the southern part of Nghe An province (rapidly increased from over 80% in early 1990s to over 90% in the early years of the 21^{st} Century). Not surprisingly, many urban areas in the country have comparatively high rates of literacy among people aged 15 year and over, at least when compared to surrounding rural neighborhoods as in Hanoi, Ho Chi Minh City, etc..

In short, the spatial patterns in education facilities, literates and enrolments were quite obvious in Vietnam. As we have discussed above the differences among the regions were rather sharp in some criteria such as enrolment rate and literates, but a bit lighter in education facilities. The latter occurs because the government of Vietnam has put efforts towards achieving equal education to all regions by improving education facilities in upland and backward areas. Despite the government's efforts, literate and enrolments rates were still much lower in the uplands compared to the lowlands. This is because demography in the upland areas is characterised mostly by ethnic minorities, whose customs, cultivation and living habits are different. They do not consider education as important. All these factors made the drop-out rate and repetition rate higher in the upland areas than in the lowland areas. Therefore, the education completion rate was lower in the uplands than in the lowlands. In addition, most indicators with respect to education were better in the North than in the South. Education enrolment and literacy rate were higher in the North compared with the southern part of the country (see Table 5.3 and Table 5.4). This can be explained as that traditionally education has been given much more attention by the households in the North than in the South, where the natural conditions are very much favorable to agricultural activities that make their livelihoods much easier than in the North (World Bank Vietnam, various years; United Nations in Vietnam, 2003; Vietnam Academy of Social Sciences, 2006; National Centre of Social Sciences and Humanities of Vietnam, 2001).

5.4. Description of the data and the choice of indicators

It is a fact that education enrolments and literacy rely heavily on the education facilities available. Education facilities are referred to as investments made in education while education records like enrolment rates or literacy rate are considered as returns from such investments. Therefore, there is a close link between the two as investment in education in a region may boost up enrolments and literacy rate in that region. Thus, the analyses of regional inequality for education would be of more importance if we analyze education facilities and enrolments separately, and then explore the relationship between education facilities and enrolments in order to pinpoint the impacts that education facilities may have on enrolments and vice versa.

We adhere to the principle of indicator selection process as that the chosen indicators must well capture a majority of dimensions in education facilities and enrolment rates that are available in the original data. In addition, the indicators selected should be well presented. Furthermore, these indicators would primarily intend to support regional educational authorities and the Ministry of Education and Training in monitoring progress in improving and maintaining the education records.

In addition, it is worthy to note hereby that higher education will not be taken into consideration, because in Vietnam all the higher education institutions have been established in the main metropolitan cities like Hanoi, Ho Chi Minh City, Da Nang, and Hai Phong. Students, who want to seek for higher education, have to travel from their place of residence to these cities. Thus, if we include higher education into the analysis, we could arrive at biased findings about the distribution of education facilities among regions. However, we are aware that there is a functional relationship between the proportion of students by regions entering higher education and the quality of lower education of that region. Based on this functional relationship we could explore the impact of lower education quality (facilities for lower education) on the number of students entering higher education. Thus, due to a lack of systematic data on proportion of students entering higher education by regions for the study years, we opt to rule out this aspect in this thesis.

Under these considerations, with reference to the availability of data, two sets of indicators will be selected for the analyses of regional education inequality in education facilities and enrolments. Education facilities are referred to as inputs for education or investments made in education while the enrolments are considered as outputs or outcomes of such investments. Therefore, the two separate analyses for input and then for output would help us to go a step further in comparing the two resulting distributions namely spatial distribution for the input-based composite index and another for the output-based composite index.

We organize this section as follows. Section 5.4.1 will elaborate the description of the data used for the analysis and the definitions of indicators. Section 5.4.2 presents two sets of indicators for education facilities and education enrolments respectively.

5.4.1. Description of data and definitions of indicators

To measure and analyze multidimensional regional inequality for education records and education facilities, we need data which contain the various dimensions of education representing education enrolment records and education facilities at the provincial level. Education enrollments records refer to the number of students by provinces that enter each education level recorded at the study years. Education facilities include numbers of teachers by provinces at each level of education and numbers of schools by provinces for each education level recorded in the study years.

Education facilities considered in this study include education teachers and education institutions. Education teacher is defined as the number of teachers at an educational level (either primary, lower secondary or upper secondary education) accounted for 1000 pupils/students at the corresponding education level while education institution is defined as the number of education institutions at an educational level (either primary or lower secondary or upper secondary education) accounted for 1000 pupils/students at the corresponding education level. The detailed definition for each variable is provided later in the section on the choice of indicators.

The data used for analysis are provided by the GSO (General Statistical Office of Vietnam). Information concerning the education facilities and enrolments are gathered from the various issues of Statistical Yearbooks. GSO used the data provided by the department of statistics at provincial levels, which was reported by the office of statistics at district levels. In addition to the data concerning education facilities, we used population data which were also provided by the same channels from GSO, at provincial and district levels. The population data were used for defining the education indicators.

Some information relating to education facilities and enrolments were gathered from the official publications of United Nations, Vietnam office, and from the office of World Bank.

Due to the practice that the data were collected from several sources, the problem relating to the quality of the data should be elaborated and discussed herewith. Despite gathering from several sources, the data are reliable and its quality is warranted because there was very close coordination and cooperation between GSO of Vietnam, the office of World Bank Vietnam and the office of UN in data gathering methods and definition of indicators. The data are usually used commonly among these organizations in Vietnam.

5.4.2. The choice of indicators

Following the discussion above and in view of the availability of the data, the following two sets of indicators were selected for the analysis of regional inequality in education facilities and enrolments.

a. Indicators for education facility analysis

1. Primary education institutions: the number of primary education institutions per 1000 primary education pupils at the time of study,
2. Primary education teachers: the number of primary education teachers per 1000 primary education pupils at the time of study,
3. Lower secondary education institutions: the number of lower education institutions per 1000 lower secondary pupils population at the time of study,
4. Lower secondary education teachers: the number of lower secondary education teachers per 1000 lower secondary pupils at the time of study,
5. Upper secondary education institutions: the number of upper secondary education institutions per 1000 upper secondary students at the time of study, and
6. Upper secondary education teachers: the number of upper secondary education teachers per 1000 upper secondary students at the time of study,

b. Indicators for education enrolments analysis

In Vietnam, the primary education is compulsory and the government policy sets towards universalization of the lower secondary education. However, the mandatory in primary education is not so strict. Children from poor, ethnic minority families and from backward and remote areas usually leave school for various reasons, such as lack of labor, or their parents are too poor to afford their children's education, and due to geographical constraints. Thus, they leave school in order to work in the field in order to earn more income to support their families, or relieving financial burden for their parents. In fact, in the rural areas and mountainous regions in the North, in the Central Highlands and in the remote areas in the South, annually 15-25% pupils had to leave their schools before they finished their school year at each level (GSO, 2004 and Nguyen, 2001). In addition, the economic crisis in 1997-98 did also cause the decline in the enrolment rates among regions. The indicators for the enrolment inequality analysis are the following:

1. Primary enrolment rate in primary education[44]: the number of children of official primary school age[45] who are enrolled in primary education as a percentage of the total of children of the official school age population,
2. Lower secondary enrolment rate: the proportion of student of official lower-secondary school age (from 11 to 15 years of age) enrolled in the lower secondary education to the people in the population in that age group in the year of study, and
3. Upper secondary enrolment rate: the proportion of students of official upper-secondary school age (from 15 to 18 years of age) enrolled in the upper-secondary education to the people in the population in that age group in the year of study.

One of the purposes of the study is to rank provinces in terms of a composite index with respect to the education dimensions under consideration in order to classify provinces based on the value of composite index. Provinces having a higher composite index are considered as more favorable than those having a smaller composite index. All indicators should have been transformed to have a positive sign, in order to have a positive impact on the composite index of corresponding dimensions of education.

In Vietnam, each province is mandated to govern and control all the education activities within the province. The people who register for his/her residence in a province are entitled to go to school in that province only. Children are not allowed to seek education in other provinces rather in the one they live, even in sparsely populated regions. In fact, there is a minor number of pupils still going to other provinces seeking for their education. However, seeking education in other provinces instead of the one they are living in is rarely occurring. Spillover effects in education facilities are therefore ruled out in the analysis of regional inequality in education facilities.

5.5. Results

The organization of this section is as follows. Section 5.1 presents the empirical results for the education facilities, which is followed by the analysis of empirical results for education enrolments in section 5.2. In section 5.3, we explore the relationship between education facilities and enrolments by the mapping method.

5.5.1. Results for education facilities

For the education facilities we assume that each province is considered to be a single and independent entity. Thus, education facilities in one province are available to students/children permanently residing (residence registration) in that province due to mandatory restrictions.

This section is structured as follows. The analysis of PCPC is given in section 5.5.1.1, which is followed by the analysis of regional inequality in education facilities in section 5.5.1.2.

[44]Primary education is defined as programs normally designed on a unit or project basis to give pupils a sound basic education in reading, writing and mathematics along with an elementary understanding of other subjects such as history, geography, natural sciences, social sciences, art and music.

[45]Official primary school age in Vietnam is set at six years of age, at which children are allowed to enroll into primary education.

5.5.1.1. The Partial Common Principal Component (PCPC) analysis

We follow step 1 to step 3 presented in section five of chapter three to estimate the PCPC model for all three periods in order to test the null hypothesis of the common principal component for all three periods against the alternative that the component of these periods is not the same. The purpose of the PCPC model estimate and the test is to derive a common weight for all relevant periods. The test results are displayed in Table 5.5 (column 2). With this very high test chi square (test chi-square is 88.346 with 10 degrees of freedom) and p-value is zero, the null hypothesis of a common component is rejected ($p < 0.01$). We therefore apply step 6 (section five of chapter three) to test whether the pairs: 1998 and 2002; 1998 and 2004; and 2002 and 2004 have a common principal component each.

The test results for common principal component of each pair of periods are also shown in Table 5.5 (from column 3 to column 5). As the test results show, the null hypothesis of the first common component for each pair of periods is rejected for all cases (chi-square values are all very high and p-values are equal to zero at all cases). We conclude that there is no common principal component for any pair of periods, thus all three periods have their own distinct component. Therefore the weights for composite index computation are not the same. They are determined from the coefficients of the first component of each period obtained using PCA method (using step 1 in section five of chapter three).

Table 5.5. Partial Common Principal Component Tests for Common Principal Component among periods

	PCPC(1) for 1998,2002 and 2004 samples	PCPC(1) for 1998 and 2002 samples	PCPC(1) for 1998 and 2004 samples	PCPC(1) for 2002 and 2004 samples
(1)	*(2)*	*(3)*	*(4)*	*(5)*
Estimated parameters	53	37	37	37
Likelihood ratio test χ^2	88.346	31.694	46.626	41.856
Degrees of freedom	10	5	5	5
P-value	0	0	0	0

Table 5.6 shows distinct component coefficients obtained using the PCA method for all the periods (follow step 8 in section five of chapter three, together with their corresponding asymptotic standard errors[46] and the relative contribution of the first component to the trace (the proportion of total variance recovered from the original variables). The Table shows only the first principal component for the three periods

[46] Asymptotic standard error of the coefficients obtained using PCA is computed using following formula (for more details please refer to Flurry 1988, page 82, 83): $s(\hat{b}_h) = \left(\frac{1}{n} \sum_{j=1; j \neq h}^{p} \frac{l_j}{(l_j - l_h)^2} \hat{b}_j \right)^{1/2}$ where the left-hand side element is the standard error of the maximum likelihood estimate coefficient ($\hat{b}_h$) of the component h, p is number of parameter estimate, l_j is the characteristic root of the component j rather than l_h characteristic root of the component h, and b_j is the estimated coefficient of j[th] indicator rather than b_h.

(different components for different periods) because we observe that most of the variation is in the first components, which have *a very high* proportion of total variance explained (85% for the 1998 period, 87% for the 2002 period and 86% for the 2004 period). As was expected, the variable's coefficients of the first component are all positive. It is worth noticing that coefficients of the first component have rather small standard errors, which implies that there is a good separation between the characteristic roots of the first and second component. The remaining components have small variance; their joint contribution to the trace is about 15% for the 1998 period, 13% for the 2002 period and 14% for the 2004 period. Thus, these findings tell us that we can be reasonably sure to lose no more than about 15% (for the 1998 period), 13% (for the 2002 period) and 14% (for the 2004 period) of the total variance if we ignore all the components except the first one. Therefore, the first principal component produced the coefficients which are to be used to compute weights for the composite index computation. Thus, there will not be a common weight for all the periods. Instead, each period has its own distinct weight.

*Table 5.6. Principal Component coefficients and their standard errors**

Variables	1998-1st PC		2002-1st PC		2004-1st PC	
Primary education institutions	0.313	(0.067)	0.384	(0.088)	0.393	(0.049)
Lower secondary institutions	0.420	(0.059)	0.438	(0.076)	0.450	(0.053)
Upper secondary institutions	0.443	(0.093)	0.370	(0.082)	0.430	(0.076)
Primary education teachers	0.513	(0.091)	0.519	(0.102)	0.518	(0.092)
Lower secondary teachers	0.472	(0.087)	0.479	(0.095)	0.383	(0.046)
Upper secondary teachers	0.467	(0.096)	0.326	(0.068)	0.209	(0.023)
Characteristic root of 1st PC	1998		2002		2004	
Characteristic roots	3.10		2.87		2.86	
Proportion of total variance	0.85		0.87		0.86	

* Standard errors are in parentheses

Looking at the estimates of the PCA models presented in Table 5.6, the variables' coefficients of the first principal component, we find that the variable primary education teachers is the most dominant in all three periods (as their coefficients values are largest at 0.513, 0.519, and 0.518 for 1998, 2002 and 2004 respectively). This means that, for all periods, this variable has largest impacts on regional inequality in education facilities distribution in Vietnam. Lower secondary teacher seems to have up and down trends as it's coefficient value lightly increased over the 1998-2002 period (from 0.472 to 0.479), but declined in the 2002-2004 period (from 0.479 to 0.383) because of the larger difference in the number of lower secondary teachers among the periods under study. For the variables education institution, as we have observed from Table 5.6 that the numbers of primary education institutions and lower education institutions have lightly increased over times. These increases would perhaps cause steady changes in the regional distribution of education facilities over the years. The values of variable lower secondary institutions have increased to the dominant levels in the year 2002 and become rather dominant in the 2002 and 2004 periods (third largest value in the 2002 period and second largest value in the 2004 period). The variable upper secondary institutions was going up and down as its coefficient decreased during the 1998-2002 period (from 0.443 to 0.370), but increased in the 2002-2004 period (from 0.370 to 0.430). The decline during the period 1998-2002 in the number of upper secondary institutions could be caused by the downturn in the economy because of the Asian economic crisis in 1997-1998, and then, it increased because of the government policy towards improving education facilities in all regions of the country. In contrast, the variable upper secondary teacher, which is quite dominant in the 1998 sample, became the least dominant variable in the years 2002 and 2004 (as its values are smallest: 0.326 for the 2002 sample and 0.209 for the 2004 sample). This means that in the years 2002 and 2004, the number of teachers in the upper secondary education has the least impact on the regional distribution of education facilities. These findings are in conformity with the recent policy towards improving education quality nationwide, Vietnam has successfully provided a rather equally number of secondary teachers to all provinces.

5.5.1.2. The analysis of multidimensional regional inequality for education facilities

Step 8 (in section five of chapter three) is applied in this section to determine weights associated to indicators and then to compute Theil's second measures. The distinct weights for each period are derived from the indicators coefficients obtained from estimating a PCA model for each period separately, because the PCPC model does not fit the data. We use equation (4.4) in chapter four to determine the distinct weights based on the estimated coefficients presented in Table 5.6. The distinct weights are displayed in Table 5.7 below.

Table 5.7. Weights associated to indicators in the principal component analysis of education facilities

Variables	Weights		
	1998	2002	2004
Primary education institutions	0.119	0.153	0.165
Lower secondary institutions	0.160	0.174	0.189
Upper secondary institutions	0.169	0.147	0.180
Primary education teachers	0.195	0.206	0.217
Lower secondary teachers	0.179	0.190	0.161
Upper secondary teachers	0.178	0.130	0.088

From these weights, the composite index for each province was computed. Composite indices and ranking orders by 61 provinces are displayed in Table 5.16. Finally, Theil's second measures for all three periods for the education facilities based on these weights are calculated and presented in Table 5.8 below. The Table shows that the changes in total inequality were rather small but in opposite directions, as we observe that the inequality declined around 2.7 % between 1998 and 2002 (down from 0.1005 in 1998 to 0.0978 in 2002), and then increased around 4.7% during next period (up from 0.1012 in 2002 to 0.1042 in 2004). The changes mean that the inequality in education facilities declined during the period between 1998 and 2002, then increased in the period between 2002 and 2004. These changes reflected the fact that during these times the government of Vietnam set the Millennium Development Goals, in which the education improvements have been one of the main focuses in order to help reducing poverty. To improve education, the government has invested more in building new schools, training more teachers for all education levels and providing better education facilities in poor provinces in order to help these provinces to catch up with the wealthier and more advanced provinces in lowland areas. However, there were differences in practice as many teachers who were trained and assigned to take a duty in the remote, backward and mountainous areas would leave their positions sometimes after a few years of assignment (World Bank, various years). In addition, the schooling equipments installed into these areas were also of very low quality. Furthermore, the years during the 1998-2002 periods were the beginning years for the implementation of the country Millennium Development Goals, in which education was given more priority (United Nations, 2003; Government of Vietnam, 2005). More money has been invested for upgrading education facilities among regions in order to achieve the education goal of universalization of primary and lower secondary education. For the period 2002 to 2004, less attention was given to improving education facilities. As a result, the total inequality has initially decreased during the 1998-2002 period (from 0.1025 to 0.1012 in 2004), then increased in the 2002-2004 period (from 0.1012 to 0.1042). However, as the data reveal the changes in inequality were still very moderate with around three to four percent differences between the periods.

Table 5.8. Theil's second measures and its decomposition for education facilities

Periods	1998		2002		2004	
	Value	Share (%)	Value	Share (%)	Value	Share (%)
Between-region inequality	0.0597	58	0.0548	54	0.0613	59
Within-region inequality	0.0428	42	0.0464	46	0.0429	41
Total inequality	0.1025		0.1012		0.1042	

The decomposed components of the total inequality are also reported in Table 5.8. These components are very informative in helping us to understand about how many percents of the changes in total inequality are caused by changes in the within-group and/or between-group terms. It also helps us to understand the proportions of the between-region and within-region inequality in the total inequality. The decomposed components of the Theil's measures tell us that there were different changing patterns between the two terms over time. The between-region inequality, which makes up a larger part in total inequality declined during the 1998-2002 period (from 0.0597 to 0.0548), and then increased in the next period between 2002 and 2004 (from 0.0548 to 0.0613). In contrast, the within-region inequality has increased during 1998-2002 period (from 0.0428 to 0.0464), then declined in the 2002-2004 period (from 0.0464 to 0.0429). This decline in the within-region inequality could be caused by the unequal distribution of education facilities within a region. To some parts of a region, more teachers and investment for building education institutions were assigned, while some parts got less of those. If we look at the evolution of both components of the inequality in education facilities over time, we observe that they have increased during the 1998-2004 period. This implies that the distribution of education facilities among regions in Vietnam has been a bit more unequal over time from 1998 to 2004. For the within region inequality, the change over time was not much (from 0.428 in 1998 to 0.429 in 2004). This implies that the opportunities to access to education facilities among people within a group were almost unchanged from time to time.

Figure 5.3 presents the trends of inequality and its compositions over time. The figure shows that the two decomposed terms of the inequality were having a different trend over time. Because the sum of the two terms of the inequality is the total inequality, the decline in the within-region inequality would be offset by the increase in the between region inequality and vice versa. As the figure shows, the within-region inequality line was in the increasing trend during the 1998-2002, and then was in the deceasing trend in the period 2002-2004. Meanwhile, the between-region inequality was first in the decreasing tendency during the 1998-2002, and then it changes direction upwards during the period 2002-2004. The line to total inequality has its shape similar to the between term as it first decreased during the 1998-2002, and then increased during the 2002-2004.

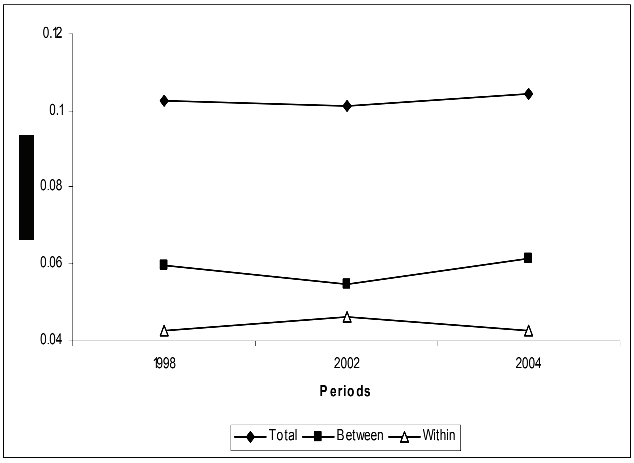


Figure 5.3. Trends of Theil's second index and its decomposed terms: the case of education facilities

Figure 5.4 below presents the clusters of regions with least favored, medium favored and most favored with respect to the composite index of the education facilities provided in Table 5.16. As we have argued earlier the higher the composite index the more favored the region is. As we observed from figure 5.4 and the statistics of the composite index provided in Table 5.16 we found that in all three periods the most parts of the North West, some parts of the North East and the Red River Delta are the most-favored regions. For the case of 1998 period, in addition to the above listed regions, some parts of the Central Highlands, the South Central Coast are also the most favored, while there is a small part in the South Central Coast most favored for the 2002 period. In the 2004 period, the situation of education facilities seems to be better because more regions are found to be most favored with respect to the composite index of the education facilities in the Central Highlands, the South Central Coast and the Mekong River Delta (Figure 5.4 and data presented in Table 5.16).

These findings were in line with the practice of investments by the government that aimed at improving education in targeted mountainous, backward areas in order to bridge education gap between lowlands and highlands stated in the government's MDGs (Millennium Development Goals). Increasing and improving education facilities in the targeted poor and backward areas should result in higher enrolments rates and higher school completion rates in these areas.

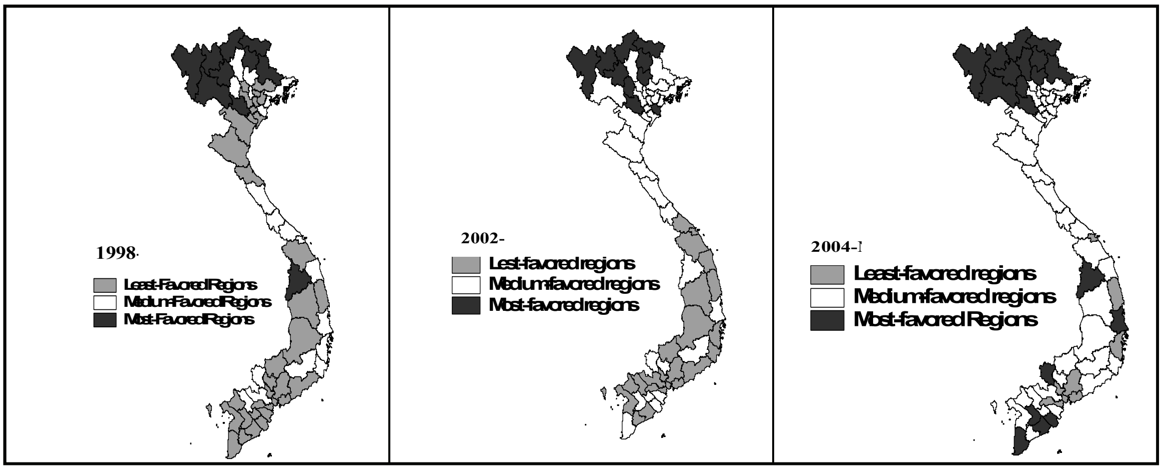


Figure 5.4. Clusters of favored regions with respect to education facilities[47]

5.5.2. Results for education enrolments

In this section, we examine regional disparities in education enrolments among provinces in Vietnam by computing and analyzing Theil's second measure and its decomposed components (between region and within region inequality) for the multidimensional regional inequality in education enrolments.

The organization of the section is as followed. The empirical results of composite index estimation with the relevant tests of common principal components and consideration of covariance matrixes of the studied samples are given in section 5.5.2.1. The analysis of spatial inequality for education enrolments is given in section 5.5.2.2.

5.5.2.1. Results of composite index estimation for education enrolment

We followed steps 1-3 (presented in the section five of chapter three) for the weights and then the composite index computation. The model estimates and its relevant tests are provided in Table 5.9 (column 2) and Table 5.10 below. The statistics provided in Table 5.9 show that the test for the null hypothesis of a common component for all the three periods against the alternate that each period has its own distinct principal component has been strongly rejected (p-value = 0, Chi square = 104.1). Then we go further to Step 6 (see section five of chapter three) to construct another test for the null hypothesis of a common component for a pair of the periods: 2002 and 2004. The test results are also given in Table 5.9 (column 3). With very high p-value *(p = 0.67),* we cannot reject the null hypothesis of a common component between the two samples: 2002 and 2004. We therefore conclude that there is a first common component for the 2002 and 2004 periods. As the null hypothesis of the test of the common principal component in which there was the inclusion of the 1998 period has been rejected, so the 1998 period has its own distinct component (first principal component obtained from using PCA method). Therefore, the indicator coefficients of the period 1998 will be the first principal component obtained using the PCA method (following step 8 in section five of chapter three), while those of the

[47] In this case, the favored region is defined as: the least-favored regions are those having the composite index ranging up to 0.259; the medium-favored regions are those having the composite index ranging from 0.259 to 0.315; and the most-favored regions are those having the composite index over 0.315.

2002 and 2004 periods are the first common component obtained using the PCPC method (following steps 1,2,3,6 and 4 in section five of chapter three).

Table 5.9. Partial Common Principal Component tests for education enrolments

	PCPC(1) for the 1998,2002 and 2004 samples	PCPC(1) for the 2002 and 2004 samples
Estimated parameters	14	10
Likelihood ratio test χ^2	104.1	0.81
Degrees of freedom	4	2
P-value	0	0.67

Table 5.10 presents the first principal component for the 1998 period and the exact maximum likelihood PCPC estimates of the common component coefficients for the next two periods: 2002 and 2004 with their corresponding standard errors and proportion of total variance recovered by the first principal component (for the 1998 period) or by the first common component (for the 2002 and 2004 periods). The calculated standard errors of the indicators coefficients were small implying a high consistency of the component coefficients. A very high proportion of total variance recovered by the first principal component (in the case of 1998 sample) and by the first common component (in the cases of 2002 and 2004 samples) indicates that the indicator coefficients presented in Table 5.16 are reasonable and robust from the statistical point of view.

*Table 5.10. Principal Component coefficients for 1998, Maximum Likelihood PCPC estimates for 2002 and 2004 and their corresponding standard errors for education enrollments**

Variables	1998-1st Principal Component		Eigenvectors (common component coefficients) for 2002 and 2004	
Primary education enrolment	0.482	(0.023)	0.462	(0.022)
Lower secondary enrolment	0.710	(0.181)	0.791	(0.021)
Upper secondary enrolment	0.513	(0.093)	0.612	(0.016)
Characteristic root of 1st PC	1998		2002	2004
Characteristic roots	0.038		0.036	0.041
Proportion of total variance	0.90		0.88	0.84

* Standard errors are in parentheses.

Table 5.11 displayed the covariance and correlation matrices between the estimated components for the 2002 and 2004 periods based on the estimates from the PCPC model for these two periods. We observe from the table that there was no substantial correlation between the common principal components occurring in both periods (all the correlation coefficients are very small). These small correlations further support our acceptance of the null hypothesis of a common principal component for the 2002 and 2004 periods.

Looking at the estimates presented in Table 5.10, the indicator coefficients, we find that for all the periods the most dominant variable was the lower secondary enrolment, whose coefficient takes the value of 0.71 for the 1998 period and of 0.791 for the next two periods of 2002 and 2004, while the least dominant variable was the primary education enrolment in all three periods. However, the value of each indicator coefficient did not differ much over the time. In addition, the difference between the value of the least dominant variables and most dominant variables was quite large in all three periods (0.791 compared to 0.462). This is in conformity with the fact that after almost 15 years of the *doi moi*, in which the education policy has considered to provide primary education (basic education) to all people regardless of gender, race, region or ethnic group and to make primary education compulsory to all children at their school age. These efforts contribute to making the rate of primary education enrolments similar between regions. It therefore becomes the least important factor that could contribute to the spatial inequality in enrolments in Vietnam. Meanwhile, the enrolment in lower secondary education has the largest dominant because the government wanted to boost the universalization of this level of education throughout the country. After finishing this level, most people enter the workforces, and a fewer number continues to go further to upper secondary education. Among school students, those who want to continue their education by going further to higher education (university or colleges) must complete their upper secondary education level (grade 12) and pass the school leaving examination. Then, they would have to attend an entrance examination in order to get admission to the higher education at the university or college level. As in Vietnam's tradition, people in most rural areas go to work after finishing lower secondary education. Most parents also encourage their children to leave school for work in order to earn more income for the household (Vietnam Development Report, 2005).

Table 5.11. Covariance (F_i)[48] and correlation (R_i)[49] matrices of the partial CPCs for 2002 and 2004[@]

Matrix for the 2002 sample				Matrix for the 2004 sample			
	1st PC	2nd PC	3rd PC		1st PC	2nd PC	3rd PC
	0.036	0.001	0.0004		0.04	0.001	0.0003
R_{02}/F_{02} =	0.016	0.003	0	R_{04}/F_{04} =	0.017	0.005	0
	0.035	0	0.002		0.065	0	0.003

[@] On and above diagonal of the matrices - variances and covariance; below the matrices' diagonal – correlations

[48] As in ordinary CPC analysis, we can compute covariance and correlation matrices between estimated principal components (see table 5.10) for the case of PCPC analysis. In the present case, the lower right (p-q)×(p-q) portion of the covariance matrices is calculated using formula: $F_i = \hat{\beta}^{(i)'} S_i \hat{\beta}^{(i)}$, where $\hat{\beta}^{(i)}$ is consistent estimate of $\beta^{(i)}$ and $\beta^{(i)'}$ is transposed of $\beta^{(i)}$. S_i is covariance matrix of indicators.

[49] Similarly, the formula to compute the correlation matrices can be written as: $R_i = \hat{\Lambda}_i^{-1/2} F_i \hat{\Lambda}_i^{-1/2}$, where $\hat{\Lambda}_i$ is consistent estimate of Λ_i, $\Lambda_i = diag(\lambda_{i1},......,\lambda_{ip})$, while F_i is the covariance matrix computed using above formula given in note number 9.

For more details about these formulas and how to construct the matrices of variances and correlations, please refer to the book by Flury (1988), p. 71 and p. 131, and relevant demonstrated examples.

5.5.2.2. The analysis of multidimensional regional inequality for education enrolments

We follow step 9 in section five of chapter three and make use of equation (4.4) in chapter four, the weights associated to indicators for the composite index computation were calculated based on the indicator coefficients obtained from using the PCA method (distinct weight for 1998 and by the PCPC method (common weight for both the years 2002 and 2004), and displayed in Table 5.12 below. Theil's second measures and its decomposed terms are computed and given in Table 5.13 and the trends are displayed in Figure 5.5 below.

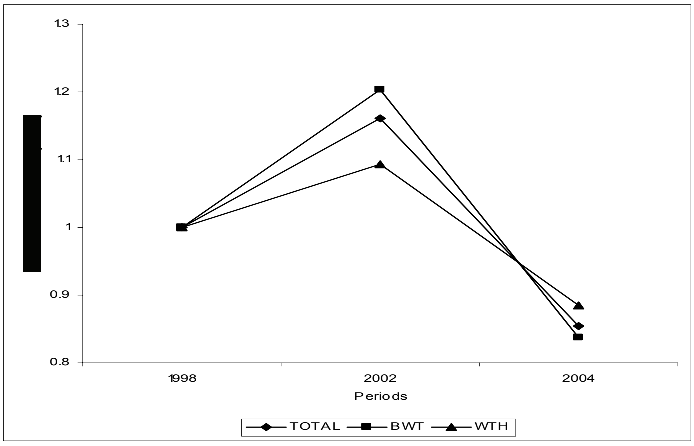
Table 5.12. Weights associated to indicators of the education enrolments

Variables	Weights	
	1998	2002 and 2004
Primary education enrolment	0.283	0.248
Lower secondary enrolment	0.416	0.424
Upper secondary enrolment	0.301	0.328

Table 5.13. Theil's second measure and its decomposition for education enrolments

Years	1998		2002		2004	
	Value	Share (%)	Value	Share (%)	Value	Share (%)
Between-region inequality	0.153	61.4	0.184	63.7	0.128	60.1
Within-region inequality	0.096	38.6	0.105	36.3	0.085	39.9
Total inequality	0.249		0.289		0.213	

As Table 5.13 reveals Theil's second measure of the total inequality in the case of education enrolments increased over the 1998-2002 period and declined in the next one (2002-2004 period). These different trends in Theil's second index between the two years can be explained. The increase in the Theil's index during the period 1998-2002 would be caused by the changes in the enrolment rate in primary education among regions in the country. The trend in inequality in enrolments was in contrast with the trend in inequality in education facilities. Despite being encouraged by the government, the number of pupils enrolled in primary education was getting more differently between upland and lowland areas (see Table 5.3). In addition, the changes in demography from year to year could also cause the large gap between enrolments among regions in the country. In the next period: 2002-2004, the inequality declined because the number of school students enrolled equally among regions. Similarly, if we look at the decomposed terms we observe that both terms (between-region and within-region inequality) have also experienced the same pattern as the total inequality did. The Theil's indices (Table 5.13) have increased during the 1998-2002 period and then declined in the 2002-2004 period. These findings imply that enrolments in the year 2004 were more equal. This means that the enrolment situation has improved over the time. The shares of between-region and within-region inequality in the total inequality tell us that in all three years the between-region term makes up a major part in the total inequality (61.4%, 63.7% and 60.1% for respective 1998, 2002 and 2004) while within-region inequality accounted for just only 38.6%, 36.5% and 39.9% for the years 1998, 2002 and 2004. These results imply that the differences in enrolments between regions were more significant to the total inequality than the differences in enrolments within regions.



***Legend*:** TOTAL indicates total inequality; BWT stands for between region inequality and WTH is within region inequality

Figure 5.5. Trends in Total, Between- and Within-Region Inequality for Enrolments

Figure 5.6 below presents the clusters of regions with least favored, medium favored and most favored with respect to composite index of the education enrolments provided in Table 5.17. As we have argued earlier the higher the composite index the more favored the region is. As we observed from the figure 5.6 and the statistics of the composite index provided in Table 5.17 we found that in all the three years the most parts of the Red River Delta, the North Central Coast, some parts of the South Central Coast and some parts of the South East are the most-favored regions (shown by the darkest regions in the map). However, more regions in the North Central Coast and the South Central Coast are more favored in the case of the 2002 and 2004 samples which indicate that there were improvements in education enrolments over time in these regions. However, as we have discussed earlier, the overall picture of enrolments was not an improvement in all three years (see Table 5.3). We also observe from the table and the map that enrolment records seem to be better in the North than in the South, as a few regions in the south (located in the South East) are considered most favored with respect to the composite index of the education enrolments. All three samples shown in the map indicate that the Central Highlands are least favored with respect to education enrolments (shown as the lightest color spectra in the map of Figure 5.6) while most parts of the North West and the Mekong River Delta are medium favored with respect to education facilities (shown as the grey regions in the map of Figure 5.6).

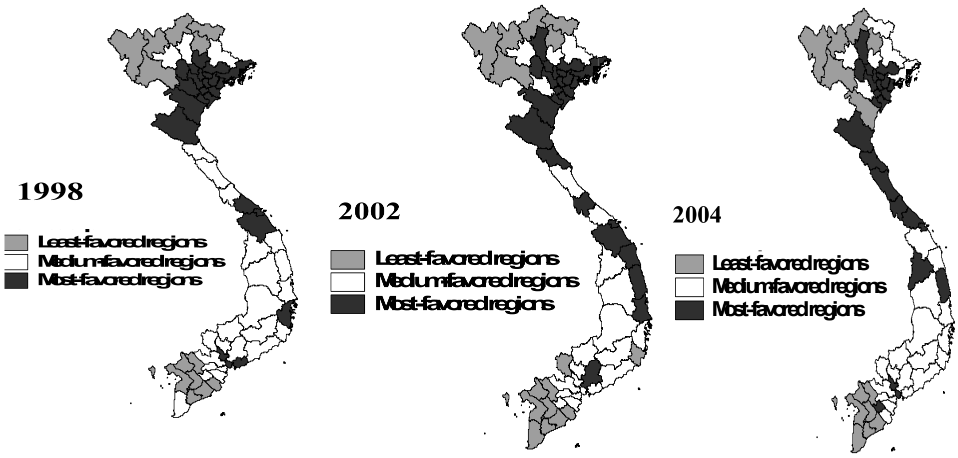


Figure 5.6. Clusters of favored regions with respect to education enrolments[50]

5.5.3. The analysis of the relationship between education facilities and education enrolments

In order to investigate the relationship between the education facilities and education enrolments by cluster analysis we use the mapping method to visualize the regions which are favorable in composite index with respect to education facilities and the regions which are favorable in composite index with respect to education enrolments in order to inspect whether the favored regions with respect to education facilities are coincided with the favored regions with respect to education enrolments or not. However, the illustration using the mapping method could not fully display and explain the correlation between the two dimensions of education as that there could be a negative or a positive relationship between investment made in education and enrolments in education as the composite index with respect to education facilities and composite index with respect to enrolments reveal that these two composite indexes did not coincide in some regions, for example for the North West where a higher composite index of the education facilities could not result in a higher composite index in enrolments.

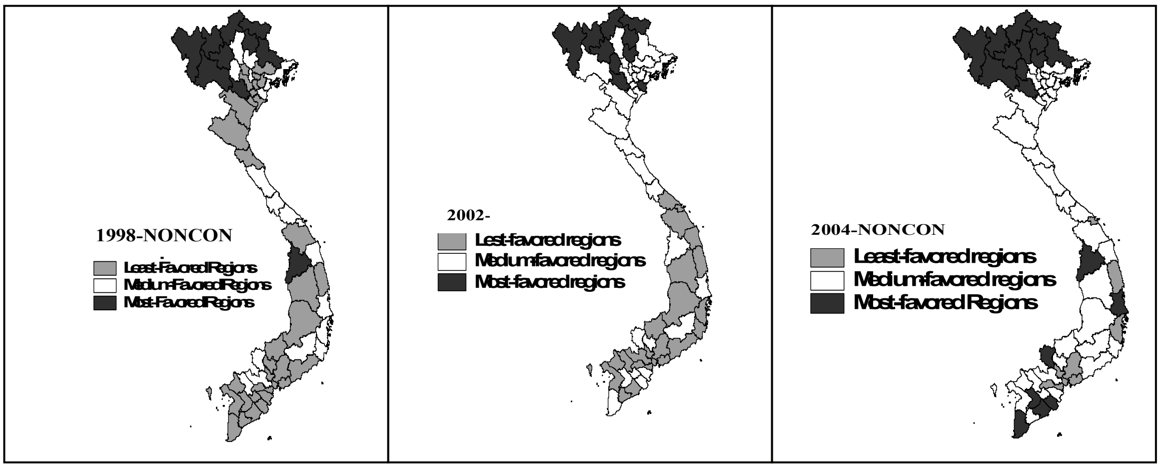
We structure this section as follows. In section 5.3.1 we compare education facilities and enrolments with each other using the visualization method based on the mapping of the composite index with respect to education facilities and composite index with respect to education enrolments in order to find out whether the favored regions with respect to education facilities are coincided with the favored regions with respect to education enrolments or not.

Figure 5.7 displays the clusters of favored regions with respect to education facilities and education enrolments over time based on the composite index displayed in Table 5.16 for education facilities and in Table 5.17 for education enrolments. Looking at both the figure and the statistics of the composite index we find that for education enrolments the most-favored regions include most parts of the low land areas in the Red River Delta (Hanoi, Hai Phong, Ha Nam, Thai Binh, Hai Duong, Hung Yen and Nam Dinh etc.), and the North Central Cost (Nghe An, Ha Tinh) and some provinces in the South

[50] For education enrolment, the least-favored regions are those having a composite index up to 0.452; the medium-favored regions are those having the composite index ranging from 0.452 to 0.61; and the most-favored regions are those having the composite index beyond 0.61.

Central Coast (Da Nang), and Ho Chi Minh city. The least-favored regions can be found in most parts of the North West, some parts of the North East and the Mekong River Delta. As the cluster analysis reveals, it is interesting to point out that the most - favored regions in education facilities turned out to be the least or medium favored regions in education enrolments (dark areas in the education facilities did not completely coincide with the dark areas in education enrolments - see Figure 5.7). This can be explained as follows. As we have discussed earlier, the *doi moi* policy gave priority to education by investing more to improve education facilities in remote and backward areas in order to boost up education in the areas and considered it as a significant measure to escape poverty. However, despite these investments, enrolment records in these regions were still low because of high drop-out rate due to poverty, people's backward cognition and other reasons such as geographical conditions that hinder children in the remote and backward and minority communities from attending schools (GSO and United Nations, 2004). These findings further support our discussion earlier that the favored in term of education facilities have not necessarily resulted in the favored in term of education enrolments in Vietnam as the higher composite index in education faculties did not coincide with the higher composite index with respect to education enrolments (as discussed earlier in the introduction section and in the section on the empirical analysis).

+ Education Facilities



+ Education Enrolments

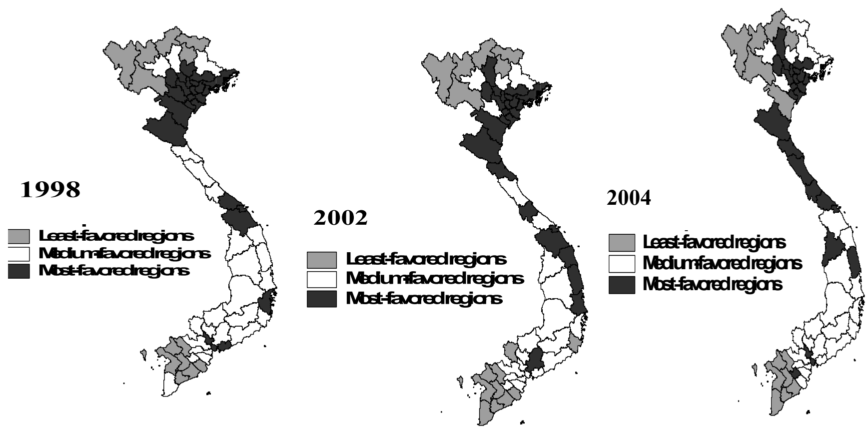


Figure 5.7. Comparisons of the clusters of favored regions with respect to the composite indices of education facilities and enrolments

Table 5.16. The composite indices for education facilities and their ranking order† by provinces

Provinces	Composite indices and ranking orders		
	1998	2002	2004
Hanoi	0.291 (15)	0.282 (20)	0.273 (36)
Hai Phong	0.285 (18)	0.238 (19)	0.285 (29)
Vinh Phuc	0.227 (41)	0.280 (21)	0.285 (29)
Ha Tay	0.259 (26)	0.300 (11)	0.302 (22)
Bac Ninh	0.243 (36)	0.263 (28)	0.269 (37)
Hai Duong	0.238 (38)	0.277 (23)	0.276 (34)
Hung Yen	0.225 (42)	0.261 (29)	0.281 (32)
Ha Nam	0.236 (39)	0.279 (22)	0.277 (33)
Nam Dinh	0.251 (33)	0.258 (32)	0.256 (42)
Thai Binh	0.265 (24)	0.312 (10)	0.309 (21)
Ninh Binh	0.242 (37)	0.273 (24)	0.284 (30)
Ha Giang	0.433 (1)	0.344 ((3)	0.476 (1)
Cao Bang	0.411 (2)	0.326 (6)	0.393 (5)
Lao Cai	0.402 (3)	0.334 (4)	0.402 (4)
Bac Kan	0.353 (5)	0.371 (2)	0.393 (5)
Lang Son	0.332 (8)	0.296 (13)	0.335 (16)
Tuyen Quang	0.311 (11)	0.291 (16)	0.340 (13)
Yen Bai	0.331 (9)	0.315 (8)	0.335 (16)
Thai Nguyen	0.298 (13)	0.328 (5)	0.348 (10)
Phu Tho	0.264 (25)	0.314 (9)	0.356 (9)
Bac Giang	0.259 (26)	0.269 (26)	0.302 (22)
Quang Ninh	0.271 (22)	0.290 (17)	0.313 (19)
Lai Chau	0.402 (3)	0.324 (7)	0.475 (2)
Son La	0.336 (7)	0.295 (14)	0.370 (7)
Hoa Binh	0.358 (4)	0.396 (1)	0.441 (3)
Thanh Hoa	0.245 (35)	0.259 (31)	0.292 (26)
Nghe An	0.253 (32)	0.271 (25)	0.301 (23)
Ha Tinh	0.253 (32)	0.264 (27)	0.273 (36)
Quang Binh	0.295 (14)	0.277 (23)	0.282 (31)
Quang Tri	0.315 (10)	0.287 (18)	0.317 (17)
T.Thien- Hue	0.271 (22)	0.230 (43)	0.265 (38)
Quang Nam	0.256 (29)	0.246 (35)	0.263 (39)
Quang Ngai	0.274 (21)	0.238 (40)	0.256 (42)
Binh Dinh	0.245 (35)	0.206 (49)	0.228 (48)
Phu Yen	0.307 (12)	0.294 (15)	0.343 (11)
Khanh Hoa	0.290 (16)	0.227 (44)	0.244 (45)
Kon Tum	0.350 (6)	0.298 (12)	0.384 (6)
Gia Lai	0.258 (27)	0.239 (39)	0.285 (29)
Dak Lak	0.257 (28)	0.219 (47)	0.260 (41)
Lam Dong	0.291 (15)	0.256 (33)	0.269 (37)
Ho Chi Minh	0.254 (31)	0.223 (46)	0.231 (47)
Ninh Thuan	0.286 (17)	0.240 (38)	0.282 (31)
Binh Phuoc	0.218 (44)	0.233 (41)	0.302 (22)
Tay Ninh	0.282 (19)	0.287 (18)	0.338 (14)
Binh Duong	0.253 (32)	0.271 (25)	0.300 (24)
Dong Nai	0.181 (46)	0.203 (50)	0.236 (46)
Binh Thuan	0.224 (43)	0.218 (48)	0.262 (40)
B.Ria-V. Tau	0.254 (31)	0.227 (44)	0.056 (49)
Long An	0.268 (23)	0.244 (36)	0.275 (35)

Dong Thap	0.282 (19)	0.247 (34)	0.293 (25)
An Giang	0.231 (40)	0.224 (45)	0.289 (28)
Tien Giang	0.258 (27)	0.238 (40)	0.251 (43)
Vinh Long	0.258 (27)	0.263 (28)	0.315 (18)
Ben Tre	0.243 (36)	0.256 (33)	0.291 (27)
Kien Giang	0.255 (30)	0.231 (42)	0.293 (25)
Can Tho	0.224 (43)	0.260 (30)	0.348 (10)
Tra Vinh	0.210 (45)	0.271 (25)	0.337 (15)
Soc Trang	0.225 (42)	0.239 (39)	0.341 (12)
Bac Lieu	0.249 (34)	0.244 (36)	0.312 (20)
Ca Mau	0.218(44)	0.256 (33)	0.358 (8)

† Rankings are given in parentheses.

Table 5.17. The composite indices for education enrolment and their ranking order† by provinces

Provinces	Composite indices and ranking orders		
	1998	2002	2004
Hanoi	0.800 (1)	0.734 (5)	0.811 (3)
Hai Phong	0.746 (2)	0.698 (11)	0.674 (23)
Vinh Phuc	0.711 (6)	0.661 (20)	0.714 (16)
Ha Tay	0.698 (8)	0.703 (10)	0.715 (14)
Bac Ninh	0.733 (3)	0.720 (8)	0.781 (5)
Hai Duong	0.726 (4)	0.718 (9)	0.715 (15)
Hung Yen	0.696 (9)	0.722 (7)	0.772 (6)
Ha Nam	0.691 (11)	0.681 (13)	0.725 (12)
Nam Dinh	0.715 (5)	0.676 (14)	0.720 (13)
Thai Binh	0.700 (7)	0.724 (6)	0.768 (7)
Ninth Binh	0.637 (19)	0.650 (23)	0.736 (11)
Cao Bang	0.415 (55)	0.442 (49)	0.495 (48)
Ha Giang	0.331 (60)	0.291 (60)	0.250 61()
Lao Cai	0.366 (58)	0.396 (53)	0.326 (60)
Bac Kan	0.424 (52)	0.450 (48)	0.414 (52)
Lang Son	0.503 (46)	0.512 (41)	0.537 (42)
Tuyen Quang	0.520 (41)	0.667 (18)	0.812 (2)
Yen Bai	0.556 (35)	0.550 (37)	0.588 (40)
Thai Nguyen	0.693 (10)	0.581 (34)	0.578 (41)
Phu Tho	0.673 (13)	0.671 (16)	0.742 (10)
Bac Giang	0.645 (17)	0.673 (15)	0.704 (19)
Quang Ninh	0.662 (15)	0.643 (24)	0.606 (35)
Lai Chau	0.319 (61)	0.278 (61)	0.337 (59)
Son La	0.417 (54)	0.336 (58)	0.396 (53)
Hoa Binh	0.634 (21)	0.576 (35)	0.620 (33)
Thanh Hoa	0.641 (18)	0.665 (19)	0.389 (54)
Nghe An	0.664 (14)	0.770 (2)	0.763 (8)
Ha Tinh	0.546 (36)	0.772 (1)	0.847 (1)
Quang Binh	0.513 (43)	0.622 (26)	0.755 (9)
Quang Tri	0.502 (47)	0.691 (12)	0.709 (18)
T.Thien- Hue	0.634 (22)	0.615 (28)	0.683 (21)
Da Nang	0.635 (20)	0.766 (3)	0.794 (4)
Quang Nam	0.645 (16)	0.736 (4)	0.630 (30)
Quang Ngai	0.583 (28)	0.670 (17)	0.656 (27)
Binh Dinh	0.607 (27)	0.654 (21)	0.685 (20)
Phu Yen	0.570 (32)	0.653 (22)	0.624 (31)
Khanh Hoa	0.630 (24)	0.597 (31)	0.601 (39)
Kon Tum	0.515 (42)	0.551 (36)	0.671 (24)
Gia Lai	0.579 (29)	0.490 (44)	0.523 (45)
Dak Lak	0.574 (31)	0.527 (39)	0.603 (36)
Lam Dong	0.610 (25)	0.605 (29)	0.660 (26)

Ho Chi Minh	0.690 (12)	0.621 (27)	0.713 (17)
Ninh Thuan	0.484 (48)	0.441 (50)	0.518 (47)
Binh Phuoc	0.527 (39)	0.507 (42)	0.536 (43)
Tay Ninh	0.568 (33)	0.453 (46)	0.523 (46)
Binh Duong	0.538 (37)	0.530 (38)	0.609 (34)
Dong Nai	0.607 (26)	0.631 (25)	0.649 (28)
Binh Thuan	0.513 (44)	0.512 (40)	0.601 (38)
B.Ria-V. Tau	0.633 (23)	0.590 (32)	0.641 (29)
Long An	0.536 (38)	0.505 (43)	0.603 (37)
Dong Thap	0.452 (50)	0.453 (47)	0.380 (56)
An Giang	0.419 (53)	0.388 (54)	0.465 (49)
Tien Giang	0.575 (30)	0.481 (45)	0.622 (32)
Vinh Long	0.522 (40)	0.603 (30)	0.677 (22)
Ben Tre	0.558 (34)	0.587 (33)	0.662 (25)
Kien Giang	0.400 (56)	0.360 (56)	0.456 (50)
Can Tho	0.452 (49)	0.402 (52)	0.449 (51)
Tra Vinh	0.424 (51)	0.418 (51)	0.533 (44)
Soc Trang	0.361 (59)	0.327 (59)	0.388 (55)
Bac Lieu	0.400 (57)	0.370 (55)	0.360 (58)
Ca Mau	0.503 (45)	0.339 (57)	0.377 (57)

† Ranking order is given in parentheses.

5.5.6. Conclusion and discussion

The implementation of the *doi Moi* in 1986 has changed the face of education in Vietnam. The main change was the regulatory mechanism from fully subsidized to partly subsidized in education in the country. In parallel to this change, the government of Vietnam paid great attention to improving education quality and quantity because she was aware that education was a good driver to poverty reduction. Therefore, investment in education would improve the quality of human resources, which was considered very important to the process of economic development. With all efforts brought about by the pro-education policies aiming at improving education towards universalization in education, there were big and significant changes in the panorama of education facilities and enrolments among regions and provinces in Vietnam.

The analysis based on the values of the indicators' coefficients of the education facilities reveals that, there was the changing tendency in the influence of each indicator on the regional distribution of education facilities. As the statistics in Table 5.6 show, the variable primary education teachers was the most dominant variable and upper education teachers was least dominant while other variables were changing their influences from year to year. These findings are in conformity with the practice of Vietnam, where more attention was paid to improving education quantity and quality in recent years. This is fruitful as the country has successfully provided a rather equal number of secondary teachers to all regions and provinces.

Measurements and analyses of multidimensional regional inequality for the two dimensions of education, namely education facilities and education enrolments conducted in this chapter along with the analysis of the composite rankings with respect to each dimension of education have brought about an interesting panorama of the multidimensional regional inequality in the two dimensions of education, and have drawn a picture of favored or non-favored boundaries with respect to the composite indexes of each dimension of education.

For regional inequality in education facilities, our research shows that total inequality, its between region and within region components go through an opposite direction between the periods. However, the change in the within region inequality is in contrast with the changes in total regional inequality and its between region component as the total regional inequality and its between region component in education facilities first decreased during the period 1998-2002 and then increased during the period 2002-2004 while the within region inequality increased during the 1998-2002 period then decreased during the 2002-2004 period. As our findings reveal, the distribution of education facilities among the regions in Vietnam was a bit more unequal in the course of time, the between-region component of inequality increased from 0.0597 in 1998 to 0.0613 in 2004. Meanwhile, the change of the within-region component of inequality was very moderate, from 0.0427 in 1998 to 0.0429 in 2004. This implies that the opportunities to access to education facilities among people within a region were almost unchanged. Therefore, the pro-poor policy culminated in education seemed not to work well in this period.

For regional inequality in enrolments, as research findings showed the patterns of total regional inequality and its decomposed terms were the same as all of them increased in the first period of time from 1998 to 2002, and declined in the period between 2002 and 2004. These findings imply that enrolments in the year 2004 were more equal. This means that the enrolment situation has improved over time. The structural shares of between-region and within-region inequality in the total inequality tell us that in all three years the between-region term makes up a major part in the total inequality. These imply that the differences in enrolments between regions were more significant to the total inequality than the differences in enrolments within regions.

Finally, we look at the comparative composite index with respect to education facilities and composite index with respect to enrolments. The findings showed that it is a sharp contrast with the case of education facilities, the most-favored regions with respect to enrolments were those in the lowland areas with least and medium-favored with respect to education facilities. The least-favored regions with respect to the education enrolments were those in the upland areas with very high-favored with respect to education facilities in terms of the composite index. These were explained in the main text of the analysis earlier that the high drop-out rate, high repetition rate were found in the areas with high incidence of poverty. These areas were coincided with those in the remote and backward areas in the upland regions of Vietnam. The findings coincided with the fact that the improvement in education facilities did not result in the improvement in education enrolments in many regions in Vietnam.

Our findings in this chapter have partly confirmed the effectiveness of the pro-education policy implemented in Vietnam during the *doi moi* process which aimed at enhancing education quality and quantity in mountainous and backward areas of the country in particular and in the whole country in general. By equipping more education facilities and assigning more teachers to remote and backward areas, the government of Vietnam has partly succeeded in promoting education in marginal and poor regions in order to help those poor to get rid of poverty. However, the effectiveness of the policy is not the case for all regions throughout the country as the research found there were many backward and remote areas in the upland regions in the country could not achieve the aims by the government. High rate of literacy, high rate of schooling drop-out, high rate of repetition as well as low education attainment among people still play a big role in hindering growth and development in remote upland areas of Vietnam.

Chapter 6
Regional Income Inequality and Regional Economic Growth

6.1. Introduction

The subject of regional income inequality in general, regional income inequality within one nation in particular and its linkage to regional economic growth has attracted much attention recently. For income inequality, the increased interest in the level, cause and evolution of global income inequality has been considerable in the 1990s. This has resulted in an extensive literature emerging in recent years that has focused on measurement, decomposition and on the study of changes of the distribution of income within a country that has regionally developed over time. The increased interest in income inequality is a result of an increased awareness about income inequality, poverty and their linkage to non-income inequalities and the increased the availability of income distribution data. With regards to a country in the process of economic transformation like Vietnam, researches on regional income and non-income inequality would be very important for development policy.

As discussed fairly detailed in chapter two, since the implementation of the economic reform the regional welfare inequality in Vietnam has grown. In chapter four and chapter five, analyses on regional multidimensional inequality in health and education respectively have been discussed. However, in these two chapters the regional inequality has been analyzed for a short period of non-contiguous time from 1998 to 2006 with three separate samples of data: 1998, 2002 and 2004. Besides the evolution of the regional inequality in health and education, per capita income also shows its disparities among provinces and regions in the country (World Bank, 2006). As the statistics of the provincial growth rate reveal, regional income in Vietnam is diverging as the rich provinces recorded a faster rate of economic growth than the poor ones did. It is in direct contrast with the inequality trend found in most developed countries like the United States and European Union, where the poor regions had a higher rate of growth than the rich ones (Barro, 1991 and Fingleton, 2003). In Vietnam, the rich provinces still enjoy a faster rate of economic growth than the poor ones do that is in line with the findings by Kuznets in his well-known work on the link between inequality and growth (Kuznets, 1955).

Internationally, many studies found both a negative and a positive relationship between income inequality and economic growth (Perotti, 1996; Li and Zou, 1998); and Forbes, 2000; and Barro, 2000). However, until now no systematic work has been conducted on the issue for Vietnam. This study is set to test whether there is positive or negative relationship between regional income inequality and regional economic growth rate. Thus, this study is almost the very first one to explore regional income inequality using Theil's index and then analyze the correlation between regional growth and regional inequality of income.

At present, Vietnam is at the initial stage of its economic development since the country implemented the comprehensive *doi moi* program as its economy transformed from a centrally planned to a market-oriented economy. Various explanatory studies by the World Bank and the General Statistical Office of Vietnam (GSO)(various years) make it is clear that the regional income in the country is diverging over the time as economy is growing faster. The statistics provided in the statistical yearbooks by the GSO reveal that rate of economic growth of wealthier regions is higher than that of poorer regions. In addition, regarding provincial income data the studies by the World Bank (2005) found that there is a trend of regional differences in per capita income in Vietnam and it implicitly reflects the potential causal correlation between disparities in regional per capita income and the regional rate of economic growth.

In chapter six, we examine the magnitude and trend of the disparities of the regional per capita income in Vietnam. measurement and analyses of regional per capita income inequality is in order to explore its evolution in the nearly two decades after the start of *doi moi* in 1986 using a panel of Gross Regional Product (GRP) per capita of 61 Vietnamese administrative units (provinces/cities) over the panel of years from 1990 to 2006. Then, the study proceeds to focus on examining the relationship between inequality and growth. This is an interesting task, because whereas cross-country study conducted by Perotti (1996) found a negative relationship between inequality and growth, more recent work by, among others, Li and Zou (1998), and Forbes (2000) proved that panel data reveal a positive relationship between inequality and growth. In this chapter we will attempt to find evidence for this relationship in the case of Vietnam.

Chapter six is organized as follows. Following this introduction, a short review of the literature on the issues is given in section 6.2. Then an exploratory analysis of regional income inequality in Vietnam and the discussion of methodology are structured in section 6.3 and section 6.4 respectively. In section 6.5, we develop a framework to measure regional per capita income inequality and a model for analyzing the relationship between regional economic growth and regional per capita income inequality. Section 6.6 contains empirical results. Finally, we end this chapter with the conclusions and discussion in section 6.7.

6.2. A brief review of the literature on regional income inequality and the link between income inequality and economic growth

A substantial number of economic development studies have been carried out to measure income inequality in general and regional income inequality in particular, and analyze the relationship between income inequality and economic growth both theoretically and empirically. In this section, we elaborate a brief review of the theoretical and empirical studies that focused on income distribution and the relationship between income inequality and economic growth.

In Vietnam, most empirical studies on the issue of income inequality were conducted by World Bank and GSO (General Statistical Office of Vietnam) (World Bank, 1993, 1999, 2003, 2005, 2007), based on the various Vietnam Living Standard Surveys in 1992, 1998, 2002, 2004 and 2006. These studies found the increase in income inequality, the Gini index has increased over the years (Gini coefficient rose to 0.432 in 2006 from 0.345 in 1990). These studies by the World Bank and GSO also explored income inequality in rural and coastal of Vietnam and found that there was also an increasing trend in income inequality in these areas. However, there is still a lack of systematic studies that focus on regional income inequality. Internationally, there are large amount of researches on regional income inequality. Starting from neoclassical growth theory, it is assumed that poorer regions display a higher marginal productivity of capital and should growth faster than rich ones. Barro and Sala-i-Martin (1988) referred to this process as beta-convergence and expressed that beta-convergence should be more likely to occur for regions within one country due to similarity of technology, preferences and institutions than for a cross-country sample. Another aspect is the potentially higher mobility of factors across regions than across country[51]. There is a large amount of literature discussing the economic convergence of Chinese and Russian regions, and most of these studies argue in favor of convergence of two steady states. Using neoclassical growth models, Solanko (2003), Ahrend (2002), Chen and Fleisher (1996), Cai *et al.* (2002), Demurger *et al.* (2002), and Weeks and Yao (2003) have found a clear evidence of conditional beta-convergence across Russian and Chinese regions, respectively. But, at the same time, Solanko (2003) and Bradshaw and Vartapetov (2003) found a growing dispersion of per capita income among regions in Russia. Ducan and Tian (1999), Cai *et al.* (2002), and Weeks and Yao (2003) also found the same results for the case of Chinese regions. At the same study, Weeks and Yao (2003) found evidence to reject the hypothesis that interior and coastal Chinese provinces displayed the same rate of technical progress and, subsequently, both groups are converging to distinct steady states[52].

For the relationship between regional income inequality and the regional economic growth rate, all the theoretical and empirical studies on the issue have stemmed from the seminal work by Kuznets (1955) and by Kaldor (1957), who suggested a tradeoff between equity and growth, and found that there was an inverse relationship between income inequality and economic growth. However, there were two opposite findings relating to the issue: positive and negative relationship between economic growth and the level of per capita income inequality. Recent theoretical literature often predicts a negative link between the two. Among those studies, we find studies by Galor and Zeira (1993), Alesina and Rodrik (1994), Person and Tabellini (1994) that found support for this prediction. Empirically, most cross-country studies using a panel of international data by, among others, Alesina and Rodrik (1994), Persson and Tabellini (1994), Perotti (1996), and Easterly (2001) find support for a negative relationship between income inequality and economic growth. However, Forbes (2000) suggested that country-specific, omitted variables were the cause of a significant negative bias in the estimations of the effects of inequality on growth. He concluded that fixed effects estimations yield consistent results of a positive short-term correlation between inequality and growth. Another study by Barro (2000), in which he used a larger sample and three-stage least squares method, had found a positive relationship between inequality and growth in developed countries and a negative relationship between inequality and growth in developing countries. Another finding on the relationship between inequality and growth by Banerjee and Duflo (1999) insisted that the relationship between growth and inequality is non-linear and changes in inequality are usually associated with lower subsequent growth.

For an economy of transition like Vietnam, there was evidence of relationship between regional income inequality and economic growth in China that is contradictory to the findings by Alesina and Rodrik (1994), Persson and Tabellini (1994), and Perotti (1996). In their study using an augmented Solow growth model with cross section and panel data for the period from 1978 to 1993, Chen and Fleisher (1996) have found

[51] Several papers mention the existence of factor mobility limiting institutions: Berkowits and Dejong (2001) discussed trade barriers between Russian regions; Zhai *et al*. (2003) describe the labour mobility limiting effect of the Chinese household registration system and land policy.

[52] For further analyses exploring potential determinant of regional inequality, see Tian (1999), Zhang and Kanbur (2001), Fedorov (2002) and Dolinskaya (2002). The literature dealing with Chinese regions is much larger than for Russia. Therefore, only selected studies are mentioned in this chapter. For more detailed see Demurger *et al.* (2002).

evidence of conditional convergence of per capita production across china's provinces during the studied period and projected that, in the near term, overall regional inequality as measured by the coefficient of variation is likely to decline modestly but that the coast/non-coast income differential is likely to increase somewhat. Another study on the China's economy by Xu and Zou (2000) found that the correlation between growth rate and Gini coefficient (coefficient of inequality) is positive, which seems to support the Kuznet's (1955) inverted U-curve. However, in an analysis of the causal mechanism relating inequality and growth in India and China, Quah (2002) did not find any evidence about the relationship between inequality and growth. In contrast, Cai *et al*. (2002), in their research on China's economic growth, inequality and labour market distortion using a framework of convergence in neoclassical theory of growth, have found that there is an evidence of conditional convergence in China's growth, namely, per capita GDP in the initiative year is negatively related to the growth rate and that in following years labor market distortion negatively impacts regional growth rates.

6.3. Regional income inequality in Vietnam: an exploratory analysis

Generally, income inequality is rising in Vietnam as the country transforms into a manufacturing and services based economy with GDP per capita increased at an average annual rate of more than 6% during 1990 and 2006 in real terms. Income inequality has increased considerably over the period as Gini coefficient rose to 0.432 in 2006 from 0.345 in 1990 (GSO and World Bank, 2007). Besides the increase in inequality at the national level, there is evidence of regional income inequality in the country as revealed by several studies by GSO and the World Bank (1995, 1999, 2003, 2005 and 2007), but it is still known for its low levels during the initial period of the *doi moi*. In these studies, GSO and the World Bank used the household living standard survey data in 1993, 1998, 2002, 2004 and 2006 to calculate the Gini coefficient for eight regions in the country. In real terms, Vietnam had a GDP per capita around 150 US dollars in 1984. Twenty two years later in 2006, this figure had risen to 722 US dollars, corresponding to an average growth rate of per capita GDP of around 2% per year during the late 1980s and 7-8% per year during the 1990s and 2000s. Table 6.1 presents the GDP per capita in purchasing power parity and by 1994 fixed price. The statistics in Table 6.1 tell us that during the period from 1984 to 2006, the growth of GDP per capita in the country is a steady one.

Data on per capita GDP for Vietnamese provinces provided in Appendix 6.1 indicate that during the late 1980s and early 1990s the income disparities among provinces were quite moderate. Statistics show that around half of the provinces in the country had figures above the national average and another half of the provinces had figures below the national average. However, the differences in the provincial income were somewhere around or close to the national average. Very few provinces had figures as high as two times the national average. Only the four richest provinces (cities) in term of per capita GDP: Hanoi, Ho Chi Minh City, Vung Tau and Binh Duong had a GDP per capita of more than double the national average. In early 1990s and beginning of the 2000s, with rather high economic growth, the regional disparities in GDP per capita has changed. In 1990, the level of differences is still low as the number of provinces having a GDP per capita above the national average level is one fifth (13 provinces in absolute number). The statistics also showed that the number of provinces having a GDP per capita lower than the national average level increased from 1990 to 1995 and that the number of provinces with the GDP per capita higher than the national average is 16 in 1995. However, this trend seems to be adverse in 2000 and 2006. In 2000, there were only 12 provinces with the GDP per capita higher than the national average, down from 16 in 1995. In 2006, only 11 provinces had a GDP per capita higher than the national average, down from 13 provinces or one fifth in 1990, 16 provinces in 1995 and 12 provinces in 2000, and accounted in that year for only one sixth. This evidence proves that there are increasing disparities in regional per capita GDP in Vietnam. It is to note that in 2006, among the provinces having the GDP per capita higher than the national average level, most of them were located in the southern part or in the South Central coast of the country (eight provinces). Only three provinces (Hanoi, Hai Phong and Quang Ninh) in the North have the GDP per capita higher than the national average. There was no any province in the North West, the North Central Coast and the Central highlands (the regions were considered as poorest in Vietnam) with the GDP per capita higher than the national average. These findings proved that the regional disparities in GDP per capita are widening between regions.

Table 6.1. Gross Domestic Product (GDP) per capita, in current international dollars (PPPs) and in 1994 prices*

Years	1984	1986	1988	1990	1992	1994	1996	1998	2000	2002	2004	2006
GDP per capita in PPP	682	734	809	941	1100	1301	1563	1785	2040	2369	2709	3384
GDP per capita in 1994 prices	151	172	194	216	244	289	340	375	415	440	568	722

*Source: World Bank and GSO of Vietnam. * PPP indicates purchasing power parity.*

The statistics provided in Appendix 6.1 also tells us that most provinces in the north and the north central coast regions have a level of GDP per capita very well below the national average. Statistics provided in Appendix 6.1 also prove that most provinces in the southern regions have a level of the gross provincial product (GPP) per capita close to the national average level of GDP per capita, while the level of GPP per capita of most northern provinces is lower than the national average level of GDP per capita: equivalent to from only 27% to more than half of the national average level of GDP per capita while the disparities in the GDP per capita in the most southern provinces is close to or higher than the national average level: changing from 83 percent to around four times of the national average level of the per capita GDP.

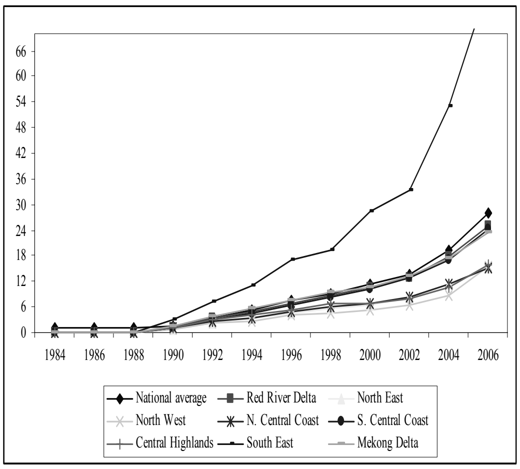
*Table 6.2. GDP per capita by eight regions in selected years**

Code	Regions	1990	1995	2000	2006
1	Red River Delta	516	2466	4578	10319
2	North East	364	1641	2709	6166
3	North West	358	1473	2212	6377
4	North Central Coast	416	1828	2863	6258
5	South Central Coast	463	1338	4161	9988
6	Central Highlands	488	1967	2761	6489
7	South East	1178	5317	11642	33365
8	Mekong River Delta	606	2737	4377	9585
National Average		*568*	*2587*	*4628*	*11571*

Source: Statistical Year Book, Various Years. The figure is measured in 1,000 Vietnamese Dong.*

Looking at the regional scenery (eight administrative regions) we observe that during the high rate of growth in the second half of 1990s and early 2000s, the evolution of per capita income by provinces in Vietnam has shown a diverging trend over time as the disparities in per capita income among regions have been widened. The statistics provided in Table 6.2 (GDP per capita by eight regions in Vietnam) help understanding the disparities in per capita income among regions in Vietnam. From the table we observe that the region with the highest level of per capita income is the South East - which is located by a cluster of high income provinces including Ho Chi Minh City, Vung Tau, and Binh Duong. The level of GDP per capita is around two times higher than the national average in 1990, but it has increased to around three times higher than the national average in 2000 and in 2006. If we compare the level of GDP per capita of the South East region (cluster of the high income provinces including Ho Chi Minh City, Binh Duong, Dong Nai and Ba Ria - Vung Tau) to the poorest region (the North West; cluster of low income provinces including Son La and Lai Chau), we find that in 1990 the GDP per capita of the South East is just 3.3 times higher than that of the North West, but this difference has increased in the course of time as in the year 2000, the GDP per capita of the South East is 5.3 times higher than that of the North West. The differences among regions also increased over time, see Table 6.2. At the regional scenery, only one region has the GDP per capita higher than the national average level, namely the Southeast. The statistics provided in Table 6.2 further consolidate our earlier findings that among the provinces that have the GDP per capita higher than the national average, most of them are in the southern part of the country. Meanwhile all regions in the North have their GDP per capita lower than the national average. This evidence is further proof for the earlier conclusion about increase in regional inequality in GDP per capita in Vietnam over the period studied.

We illustrate the tendencies of changing patterns of GDP per capita provided in Table 6.2 in Figure 6.1 below. Looking at the figure we observe that the per capita GDP of most regions in Vietnam is diverging away from each others, but still close to the level of national average except the line presents the case of the Southeast where the most prosperous and wealthy cities and provinces as Ho Chi Minh City, Binh Duong, Dong Nai and Ba Ria-Vung Tau are located. We observe that the line is well diverging away from the national average line and the level of GDP per capita levels of all remaining regions. This implies that the regional differences in per capita GDP are increasing.



Note: We have set the national average GDP per capita in 1984 equal to one and other values are adjusted accordingly.

Figure 6.1. Evolution of GDP per capita in current international dollars (in 1994 price) by eight regions

6.4. Methodology

In this section, we develop two frameworks. One is used to measure regional income inequality, section 6.4.1, and the other is used to analyze the relationship between regional inequality and regional growth in Vietnam in section 6.4.2.

6.4.1. Method to measure regional income inequality in Vietnam

In the cases of other dimensions of welfare such as education and health presented in chapter four and chapter five, a more proper approach for measuring and analyzing regional inequality is the multidimensional technique using Theil's second measure (see chapter four and chapter five). For the case of income, in direct contrast to other domains of welfare such as health and education, the multidimensional approach could not be applied because income is not considered as a multi-component indicator of welfare. Therefore, we develop a new approach to measure and analyze regional income inequality. We also use Theil's index of inequality and its decomposition method for this analysis. Like the case of measuring multidimensional regional welfare inequality developed in chapter three, we also use Theil's index in this chapter. Unlike the case considered in chapter three, however, the main variable used in this chapter is income per capita, a single dimensional indicator to be used for the analysis.

We apply the one-stage decomposition method of Theil measure of inequality developed by Anand (1983) and later used by Akita (2003) in his study about regional income inequality for China and Spain. We adapt this to the case of the Vietnam provinces. To use this method, we need to set the country in a hierarchical order. To do so, let's consider the following hierarchical structure for Vietnam which is commonly used in the regional studies in the country (in crescent order of desegregation):

Country \ Regions \ Provinces

In this structure, the country is composed of eight administrative regions (see Appendix 2.1 of chapter two). Each region has a different number of provinces. At present, Vietnam has 64 provinces (since 2005). But in this study data collected refer to the period from 1990 to 2006 when the country had only 61 provinces. Thus, we use this old structure of 61 administrative provinces for the analysis.

Using this structure, total regional income inequality (T_{TT}) can be measured by the following Theil's index using the population shares as weights:

$$T_{TT} = \sum_i \sum_j \left(\frac{N_{ij}}{N} \right) \log \left(\frac{N_{ij}/N}{Y_{ij}/Y} \right) \qquad (6.1)$$

where:

Y_{ij} is the income of province j in region i,

Y is the total income of all the provinces $\left(Y = \sum_i \sum_j Y_{ij} \right)$,

N_{ij} is the population of province j in region i, and

N is the population of all provinces $\left(N = \sum_i \sum_j N_{ij} \right)$.

As the regional inequality can be decomposed into between-regions and within-region inequality, and Theil's index poses this characteristics, so if we define T_{Bpi} as the between-provinces income inequality for region i:

$$T_{Bpi} = \sum_j \left(\frac{N_{ij}}{N_i} \right) \log \left(\frac{N_{ij}/N_i}{Y_{ij}/Y_i} \right) \tag{6.2}$$

then Theil index T_{TT} in equation (6.1) can be decomposed into:

$$\begin{aligned} T_{TT} &= \sum_i \left(\frac{N_i}{N} \right) T_{Bpi} + \sum_i \left(\frac{N_i}{N} \right) \log \left(\frac{N_i/N}{Y_i/Y} \right) \\ &= \sum_i \left(\frac{N_i}{N} \right) T_{Bpi} + T_{BTR} \\ &= T_{WTR} + T_{BTR} \end{aligned} \tag{6.3}$$

where:

Y_i is the total income of region i $(= \sum_j Y_{ij})$,

N_i is the total population of region i $(= \sum_j N_{ij})$, and

$T_{BTR} = \sum_i (\frac{N_i}{N}) \log(\frac{N_i/N}{Y_i/Y})$ measures income inequality between regions.

Equation 6.3 is the ordinary Theil inequality decomposition, in which the total income inequality (T_{TT}) is the sum of the within-region term (T_{WTR}) and the between-region term (T_{BTR}). The within-region term is a weighted average of the between-province income inequalities for each region (T_{Bpi}).

6.4.2. Method to explore the links between regional income inequality and regional growth

In this section we develop the model to explore the relationship between regional income inequality and the growth of regional income per capita by assessing and analyzing the impact of regional income inequality on regional economic growth. The starting point for this analysis stems from the work by Kuznets (1955), in which he speculates about the existence of a systematic relationship between inequality and the process of economic development. In his work, Kuznets finds that income inequality increases in the early periods of economic development, when the economy experiences the transformation from a rural and agricultural economy to the industrial one, and then it decreases when the modern technologies and structure have been brought into the entire socio-economic activities. In his work, Kuznets finds that there is the inverted U-shaped relationship between inequality and per capita income. This relationship is well-known as the Kuznets curve.

In the economic literature, there are two schools of thoughts about the relationship between inequality and growth. The first one considers inequality to be detrimental for growth while the second one predicts that the presence of an unequal distribution of resources is an important determinant for the development of an economy.

In this chapter we follow the idea developed by Kuznets and extended the framework to measure the impact of regional inequality on growth used in the study by Panizza (2002). The model considers growth as a function of initial inequality, income, human capital, public investment, age structure and regional dummy variables. This type of model is similar to the ones used in most empirical work on growth and inequality. The only change in our model from the Panizza's model is the addition of regional dummy variables. The regional dummies are added to control for possibly different growth patterns among regions. The growth model used in this chapter is the simple cross-province regressions of the following type:

$$g_i = \alpha + \beta y_i + \lambda I_i + \gamma X_i + \rho RD_i + \varepsilon_i \qquad (6.4)$$

where: g_i is the average annual growth rate of per capita GDP of province i over the period under consideration; y_i is province i's log of per capita GDP; I_i is variable capturing income inequality measured using Theil's index; X_i is a matrix of control variables associated to province i, and RD_i is a matrix of regional dummy controlling for the possibility of different growth patterns in eight different regions of Vietnam (Red River Delta, North East, North West, North Central Coast, South Central Coast, Central Highlands, South East and Mekong River Delta). All explanatory variables are measured at the beginning of the growth period analyzed.

The matrix X_i includes a set of control variables that are likely to be correlated to both the income distribution and economic growth, see Perotti (1996) and Barro (2000), and we control for the stock of human capital (the variable second measures the percentage of adults with secondary school degrees), the public investment (the variable invest measures the ratio of public investment to GDP), and age structure (the variable old measures the percentage of the population age 65 years and over).

To control for province-specific (regional-specific) effects and period-specific (time-specific) effects (or control for unobserved time-invariant state characteristics) we develop another model based on the growth model in equation 6.4. The new growth model is of the following form:

$$g_{(t,t+n),i} = \beta y_{t,i} + \kappa I_{t,i} + \gamma X_{t,i} + \alpha_i + \eta_t + \varepsilon_{t,i} \qquad (6.5)$$

where: $g_{(t,,t+n)}$ is the annual growth rate of per capita GDP from period t to $t+n$; $y_{t,i}$ is the province I's log of per capita GDP in time t; α_i is a province-specific (regional specific) intercept, and η_t is a period-specific (time-specific) intercept.

All other variables are defined as in model 6.4. Apparently, the coefficient κ in (6.5) looks similar to the λ in (6.4), but it is worth, however, to note here that the coefficient κ in equation 6.5 has a different interpretation from the λ in equation (6.4). While the former measures the correlation between changes in inequality and changes in growth within a given province, the latter should be interpreted as a measure of the relationship between inequality and growth across provinces or regions (Forbes, 2000).

To estimate model 6.4, we follow Barro (2000) and Panizza (2002) to use the tool of OLS for the panel data regression and for pooled data regression. To estimate model 6.5 we follow Forbes (2000) in considering the province-specific effects to use the method of fixed effects or random effects estimation and to test the validity of the assumptions underlying each method in order to find out which is an appropriate method of estimation. It is hereby worth note to that there are advantages of using fixed effect or random effect estimators. Fixed effects estimator uses a transformation to remove the unobserved effect prior to the estimation. Any time constant explanatory variables are removed along with the unobserved effect. Meanwhile, the random effects estimator is attractive when we think then unobserved effect is uncorrelated with the explanatory variables. For the purpose of estimating equation 6.5, the major difference between these two techniques is the information utilized to calculate the coefficients. The fixed-effects estimates are calculated from differences within each province across time, while the random-effects estimates incorporate information across individual provinces as well as across periods. In order to find out which is an appropriate method of estimation, either fixed effects or random effects, we apply Hausman specification test (Hausman, 1978). All estimations and tests are carried out by using the STATA software.

6.5. Data source

The data used in this chapter come from several sources. Data to compute Theil's measures of inequality are drawn from the per capita income at both the provincial and district (an administrative unit level that is next to and smaller than the provincial level) levels which are available from the general statistical office of Vietnam. Income is measured by real gross regional product (GRP) per capita adjusted to the 1994 prices. The per capita growth rate is gathered from the statistical yearbook published yearly by the General Statistical Office of Vietnam (GSO). The rate is based on the differences between two subsequent years. Human capital statistics are collected from the archived source provided by the Ministry of Education and Training of Vietnam Public investment data are collected from the statistical yearbook of Vietnam. Age structure statistics also come from the archived source of the GSO.

Due to the data availability, this chapter focuses on growth from 1990 to 2006. In fact, Vietnam started the *doi moi* in 1986, but there is a lack of statistical data for the years before 1990. So we could not find a reliable source of data for the years before 1990 for this study. Also, data before the *doi moi* were impossible to find, so we mainly focus the analysis in this chapter to the 1990 to 2006 period.

6.6. Results

In section 6.6, we report the results on regional income inequality and the relationship between growth and inequality in Vietnam. The section is organized as follows. In section 6.6.1 we present Theil's measures of regional income inequality and provide a detailed analysis of regional inequality. The analysis of the relationship between regional inequality and regional growth, and examination of the effects of inequality on growth is given in section 6.6.2.

6.6.1. The analyses of regional income inequality

Theil's measures and its decomposed components (between-province and within-province inequality) have been computed for every year in the period under the study using equations (6.1), (6.2) and (6.3) and presented in Table 6.3.

Looking at the statistics provided in Table 6.3 we observe that the inequality in Vietnam moderately increased over the period 1990-2006. However, there was a light decline during the years 1998 and 1999. This decline would be caused by the economic downturn in the country which happened during the Asian economic crisis in 1997-1998 which badly affected the income level of the rich provinces. The decomposition presented in columns 3 and 5 of Table 6.3 tells us that during the beginning of the 1990s the within-region inequality overpasses the between-region component as the statistics showed that from 1990 to 1993, the share of within-region inequality is of 56 -52 %. However, this trend has changed at the end of the study period when the share of within-region has declined to 46-49% while the share of its counterpart increased accordingly. Besides, the decomposed terms also tell us that there is an overall increasing trend in both within-region inequality and between-region inequalities in Vietnam. This finding is somehow in line with the finding by the study by the Lu and Wang (2002) on the evolution of China's regional inequality during the 1978-98 reform period, in which they have used per capita GDP, per capita consumption, and per capita income to calculate three indices: the coefficient of variance, the Gini coefficient, and the Theil's index and found that inter-provincial and regional inequality in China declined between 1978 and 1990, but steadily widened after 1990. However, the increase in regional inequality in Vietnam is more for the between region component than for the within region component, but the rise in total inequality is moderate.

The finding results on regional income inequality in Vietnam are in line with previous findings by empirical studies in Vietnam as regional inequality increases in the course of time by World Bank (various years) using Gini index of inequality. This finding also supports the conclusion by the World Bank (2006) saying that after the *doi moi*, regional inequality in Vietnam measure in Gini coefficients tended to increase.

Table 6.3 Evolution of inequality in Vietnam over the period 1990-2006

Years	Theil's index and its decomposition				
	T_{TT}	T_{WTR}	*Share (%)*	T_{BTR}	*Share (%)*
1990	0.1034	0.0579	0.56	0.0445	0.44
1991	0.1051	0.0578	0.55	0.0473	0.45
1992	0.1088	0.0588	0.54	0.0500	0.46
1993	0.1127	0.0586	0.52	0.0541	0.48
1994	0.1159	0.0580	0.50	0.0580	0.50
1995	0.1183	0.0568	0.48	0.0615	0.52
1996	0.1205	0.0566	0.47	0.0639	0.53
1997	0.1226	0.0576	0.47	0.0650	0.53
1998	0.1212	0.0558	0.46	0.0654	0.54
1999	0.1187	0.0546	0.46	0.0641	0.54
2000	0.1201	0.0576	0.48	0.0625	0.52
2001	0.1223	0.0575	0.47	0.0648	0.53
2002	0.1287	0.0631	0.49	0.0656	0.51
2003	0.1322	0.0635	0.48	0.0687	0.53
2004	0.1354	0.0650	0.48	0.0704	0.52
2005	0.1398	0.0657	0.47	0.0741	0.53
2006	0.1481	0.0696	0.47	0.0785	0.53

Table 6.4 displays the level of inequality measured by Theil's index in eight regions in Vietnam for selected years. The statistics provided in the table show that there is a rising trend (despite some declines during the period) in inequality in almost all regions in the country except for the North West and the Central Highlands. The decline in inequality in these two regions may be caused by the increase of the prices of industrial crop products such as coffee, pepper and raw rubber, as these crops are the main economic advantages of these two regions. The income of most farmers in these two regions comes from this source. So, the increase of their price would increase the income of the farmers participating in producing these crops. Thus, their income would temporary catch up with the income of other groups working from more lucrative sectors in the regions. In addition, from Table 6.4 we find that the region with the highest level of income inequality is the South East of which the Theil's index is very much higher than the one of the remaining regions in the country while the North Central Coast is recorded to have the lowest level of inequality in GDP per capita distribution. The Theil's index of the South East is around 100 times higher than that of the North Central Coast. The reason behind this high level of the Theil's index of the South East is that the region consists of very rich cities and provinces such as Ho Chi Minh City, Vung Tau, Binh Duong and Dong Nai and also very poor provinces such as Binh Phuoc, Ninh Thuan and Binh Thuan. Thus, it is understandable that there is a very high regional inequality in this region when the GDP per capita in Vung Tau is around 40 times higher than that of the poorest province in the region like Ninh Thuan and Binh Thuan. Except for this region, the differences in Theil's index in the remaining regions are not that high.

Table 6.4. Internal regional Theil's index for selected years

	1990	1994	1998	2002	2006
Red River Delta	0.0884	0.1074	0.1194	0.1267	0.1372
North East	0.0556	0.0288	0.448	0.617	0.737
North West	0.0094	0.0054	0.0028	0.0012	0.0064
North Central Coast	0.0035	0.0030	0.0049	0.0032	0.0089
South Central Coast	0.0565	0.0717	0.0686	0.0912	0.0838
Central Highlands	0.0573	0.0151	0.0117	0.0079	0.0081
South East	0.4974	0.5259	0.3871	0.5621	0.7464
Mekong River Delta	0.0105	0.0094	0.0058	0.0083	0.011

Figure 6.2 illustrates the contribution given by the within-region and between-region components of inequality to the total regional inequality in Vietnam over the period 1990-2006, and Figure 6.3 shows the evolution of inequality and its decomposed components over the same period. Both figures are just a mean to further support the trend of inequality found and presented in Table 6.3 and Table 6.4. We observe from the figures that the Theil's index steadily increases during the 1990-2006 period (Figure 6.4). Its decomposed components presented in Figure 6.2 show the changes of between-region and within-region inequality during the period which indicates that between-region inequality usually accounts for the larger part compared with its counterpart's within-region component in the total regional inequality.

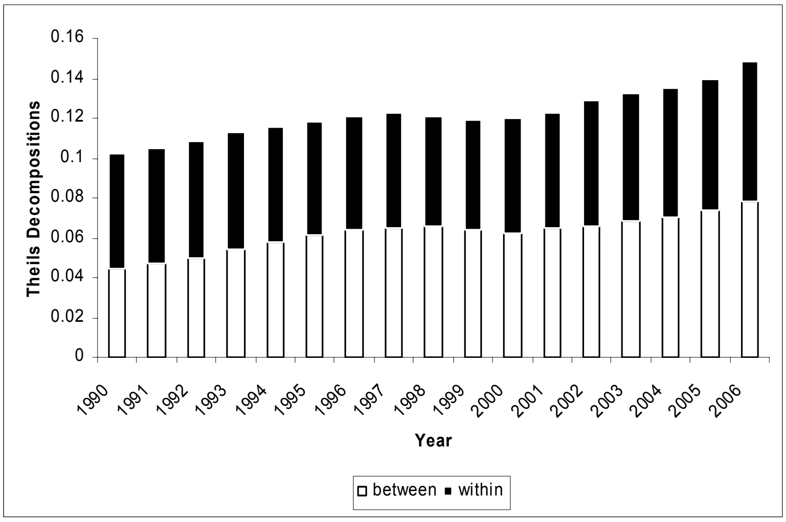


Figure 6.2. Decomposition of inequality in Vietnam over the period: 1990-2006

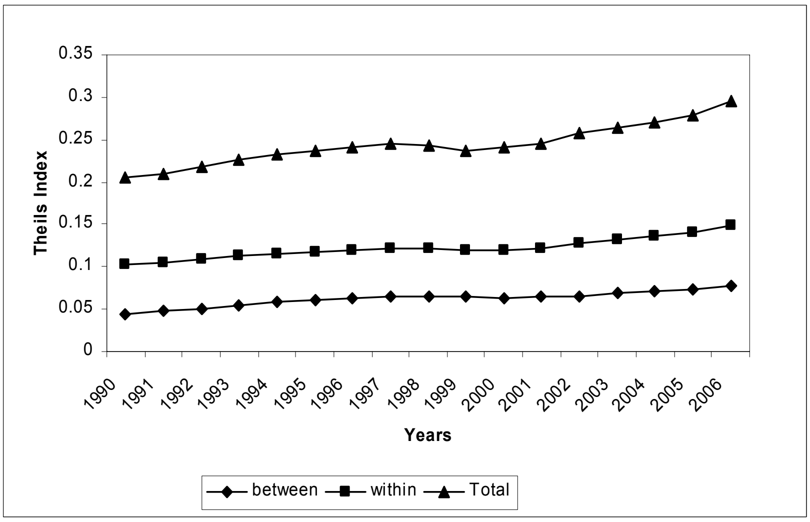


Figure 6.3. Evolution of inequality in Vietnam: overall inequality, between province inequality and within province inequality: period 1990-2006

6.7. The analysis of the link between regional income inequality and regional growth

As mentioned above, to explore and analyze the relationship between economic growth and regional inequality there are a number of estimation techniques to be used to estimate equation 6.4 and equation 6.5. We use panel data OLS and the pooled OLS technique for equation 6.4 and fixed effects or random effects technique for equation 6.5. For the latter, we report both the fixed effects and random effects estimation results and the specification test is reported later.

We start by estimating Equation 6.4 using ordinary OLS regression for the entire growth period and the results are reported in Table 6.5. Table 6.5 shows that the coefficients of the log of per capita income are negative and the coefficients of Theil's index are positive for all the three cases: without both controls and regional dummies[53]; with controls but no regional dummies; and with both controls and regional dummies. We include regional dummy into this model since our model is used for regional analysis of inequality and growth. Thus is used in order to control for effects caused by neighboring regions that need to be controlled for. However, all the coefficients of log of province per capita income are not statistically significant, but the coefficient of the Theil's index for the case of no controls is significant at 1% level while the coefficients of the Theil's index for two remaining cases are not statistically significant. Look at the coefficients of the log per capita income presented in the first row of Table 6.5 for all the three cases we found that there was a negative relationship between economic growth rate and initial level of GDP per capita in Vietnam. The statistics in row two of Table 6.5 for all the three cases tell us that there is the positive relationship between regional inequality and regional growth. Meanwhile the results reported in column three and column four of Table 6.5 for the remaining two cases: with the control variables but no regional dummies, and with both the control and regional dummy variables respectively, indicate that the positive relationship between regional inequality and regional growth in Vietnam is not significant. The three control variables included in the model indicate that public investment and human capita play a significant role in the economic growth rate of a region as their coefficients are positive and significant for public investment but not significant for human capital (in both cases: with controls and with both controls and regional dummies). Meanwhile, the share of the elderly has a significantly negative impact on the growth rate (its coefficient is negative and significant at 10% level). Also, based on the R^2 provided in the last row of Table 6.5 we find that, among all the cases, the models with the inclusion of the controls and regional dummies are better in explaining the variance than the model without controls and regional dummies (R^2 equals 11%, 26% and 28% for respective models with no controls, with controls and with both controls and dummies), but the model with both the controls and dummies give the best explanation.

Table 6.5. Panel data regression for growth rate in the 1990-2006 period

Independent variables	No Controls	Controls	Controls and Regional Dummies
Log of province's income per capita	-0.6471	-0.4165	-0.4077
	(0.485)	(0.464)	(0.501)
Inequality (Theil's index)	4.1383**	1.9482	1.9555
	(1.598)	(1.643)	(1.665)
Ratio of public investment in GDP		4.2401***	4.2172**
		(1.492)	(1.577)
Share of people with secondary education (human capital)		0.0006	0.0006
		(0.002)	(0.002)
Share of the elderly		-2.6989 *	-2.7338*
		(1.503)	(1.676)
Regional Dummy	-	-	-0.0049
			(0.099)
Number of observations	60	60	60
R^2	0.11	0.26	0.28

[53] In our study, we analyze growth and inequality for eight regions in Vietnam using regional partition that popularly used in Vietnam. So the regional dummy is defined as follows. The dummy takes the value of 1 if region is the Red River Delta; 2 if region is the North East; 3 if region is the North West; 4 if region is the North Central Coast; 5 if region is the South Central Coast; 6 if region is the Central Highlands; 7 if region is the South East; and 8 if region is the Mekong River Delta

Notes: + Standard errors are in parentheses. *** indicates a parameter which is highly statistically significant; ** statistically significant at 1% level; * statistically significant at 10% level. + The dependent variable is the growth rate of real per capita GDP. The growth rate is average for entire the sample from 1990 to 2006. All the explanatory variables are measured at the beginning of the period under study. For the variable inequality, we use raw Theil's indices computed for individual provinces based on income data at district and provincial levels. + Estimation by OLS technique.

In order to consolidate the findings from the results of estimating Model 6.4 for the entire growth episode from 1990 to 2006 (as given in Table 6.5) and to investigate whether there may be possibly different patterns of the inequality-growth relationship for different lengths of growth periods, we further explore the relationship between per capita income inequality and economic growth for different shorter growth periods following the method used by Panizza (2002). To do so we breakdown the entire growth period going from 1990 to 2006 into some growth periods of smaller length (five and ten-year growth episodes) and again to estimate model 6.4 for all these broken-down growth periods of five and ten-year periods for only one case: with both the controls and regional dummy by the mean of panel data regression. The reason behinds estimating Model 6.4 for only the case with both the controls and regional dummy is based on our observation from Table 6.5 that among all the three cases, the model with the inclusion of these two sets of variables (controlled variables and dummy) is best explained and best fit because (among three estimations) this estimation yields highest R^2 (of 0.28). The results from estimating equation 6.4 for five and ten-year growth episodes are reported in Table 6.6. In details, for five-year growth length, we have 3 regressions such as 1) for the growth period from 1990 to 1995; 2) for the growth period from 1996 to 2000; and 3) for the growth period from 2001 to 2006. For ten-year growth length, we have 2 regressions such as 1) for the growth period from 1990 to 2000; and 2) for the growth period from 1996 to 2006. In Table 6.6, we report estimate results for these 5 regressions and reports only coefficients and standard errors attached to inequality variable: Theil's index. Statistics in Table 6.6 reveal that the coefficient of Theil's index is positive in all the regressions for the two lengths of growth episodes. This finding again confirms our conclusion previously based on the results obtained by estimating model 6.4 with the entire growth period that there is positive relationship between regional inequality and growth but the relationship is significant when growth is measured over a five-year period which starting from 2001, and when growth is measured over a ten-year period which start from 1996 (for the sake of simplicity, in Table 6.6, the constant, controls and regional dummies are omitted from the table). For the remaining periods, the positive relationship between regional inequality and growth is not significant.

However, the estimation results reported in Column 4 of Table 6.5 for the entire growth episode, and those reported in Table 6.6 (all these estimates with both controls and regional dummies) for the five and ten-year growth episodes tell us that there exist differences between the results of the regressions for entire, five and ten-year growth episodes. In the entire growth episode, the results show that there is no significant correlation between growth and inequality, but in the five and ten-year growth episodes, there is significant correlation between growth and inequality in case of five-year length starting from 2001, and in case of ten-year length starting from 1996. This difference may be due to fact that the short-run relationship between changes in inequality and changes in growth is different from perspective long-run relationship, they could also be driven by the limited degrees of freedom in the regression in Table 6.6.

Table 6.6. Panel data regressions: five and ten-year growth episodes

	Length of Growth Episode	
	5 years	10 years
Starting Year	Theil's index	Theil's index
1990	2.445	1.956
	(1.699)	(1.894)
1996	0.663	3.071*
	(2.342)	(1.744)
2001	3.352**	
	(1.372)	

Notes: ** indicates coefficient is statistically significant at 1% level; * implies coefficient is statistically significant at 10% level. Standard errors are in parentheses.

Next, we estimate equation 6.4 using the pooled OLS regression for the entire growth period for all the three cases: without both the controls and regional dummies, with the controls but no regional dummies and with both the controls and regional dummies. The results of the pooled OLS regression are reported in Table 6.7 below (constant is omitted from the table). The results suggest that there is a negative relationship between regional growth and regional income inequality in all specifications, but the relationship was statistically significant only in the case when the controls and regional dummies were included. It was not a statistically significant relationship when both the control and regional dummy variables were omitted and in the case with inclusion of controls but omission of regional dummies. We also observe from the second row of Table 6.7 that all the coefficients of the log GDP per capita are positive and they are highly significant in the cases of omission of both the controls and dummies, and with the inclusion of controls but no regional dummies. This indicates that there is a positive relationship between growth rate and level of GDP per capita. If we look at the coefficients of the control variables in the case of with the controls and no regional dummies and in the case of with both the controls and regional dummies we observe that the latter case seems to be more conformity with the literature as there is positive relationship between investment and growth rate, and between education and growth rate. With highest yielded R^2 and lowest *AIC* and *BIC* values, the model with both controls and regional dummies is best explained and most preferred.

Finally, we estimate equation 6.5 using fixed-effects and random-effects techniques. The estimates are reported in Table 6.8 for three cases: without both controls, regional and time dummies, with controls and no regional and time dummies, and with inclusion of both controls, regional and time dummies. As indicated, the regional dummy is used to control for the possibility of different patterns in eight different regions of Vietnam (Appendix 4.1). The time dummy takes value of one if the observations belong to the 1990s decade and two otherwise. We observe from Table 6.8 that most of the coefficients are not significant in both estimates.

Table 6.7. Pooled OLS for growth rate: entire growth period

	No Controls	Controls	Controls and regional dummies
Log of province's income per capita	0.3753** (0.124)	0.7345** (0.129)	0.0376 (0.125)
Inequality (Theil's index)	-0.2324 (0.639)	-0.3720 (0.644)	-1.7226** (0.587)
Ratio of public investment in GDP		-0.0021 (0.002)	0.0002 (0.002)
Share of people with secondary education (human capital)		-0.0041** (0.007)	0.0014 (0.007)
Share of the elderly		-0.1895* (0.092)	0.0442 (0.084)
Regional dummy			0.6834** (0.044)
Number of observations	1031	1031	1031
Adjusted R^2	0.09	0.18	0.26
AIC	2527.32	2453.13	2241.70
BIC	2542.13	2482.76	2281.21

Notes: *p-value* in parentheses. ** indicates highly statistically significant; * statistically significant at 5%; - Dependent variable is log of annual growth rate; - Estimation is by pooled OLS technique

First, when we restrict both the place-specific effect (the regional dummy) and the time-specific intercept η_t (the time dummy) in equation 6.5 to be equal to zero, we find a positive and not statistically significant correlation between changes in regional income inequality (Theil's index) and regional growth in the both the estimates: fixed-effects and random-effects (columns two, three, five and six of Table 6.8). We also observe that in the case of without controls, regional and time dummies and with controls but no regional and time dummies, the coefficient, standard error and *t* statistic attached to the Theil's index of inequality is almost the same (*t* statistic is 1.33 in the former and 0.28 in the latter) in the fixed effects estimate while in the random effects estimates, all the coefficient, standard error and *t* statistic attached to the Theil's index steadily decrease in the case with control variables compared with the case without controls (*t* statistic dropped to 0.25 in the latter from 0.7 in the former). Columns four and seven of Table 6.8 show that, when the regression is augmented with both the regional and time dummies, the correlation between inequality and growth does not change much in the case of fixed effect estimate. In the case of random effects estimate, the coefficient attached to Theil's index of inequality changes dramatically to be negative. Particularly, the coefficient, *p*-value and *t* statistic attached to the Theil's index are almost unchanged in the fixed effect estimates (*t* statistic for the case with both controls and regional and time dummies is 1.33 in the fixed effect estimate) while they largely and steadily decrease in the random effects estimates (*t* statistic is -1.62).

Interestingly, both sets of regressions presented in Table 6.8 with or without control variables for both the fixed effects and random effects estimates suggest the presence of a positive relationship between inequality and growth but the relationship is not statistically significant. We also observe from Table 6.8 that, there is a change once the model is augmented with the regional and time dummies. In this case, the relationship between growth rate and regional inequality is still positive in the fixed effects estimate, but it is negative in the random effects estimate. However, the correlation between growth and inequality very slightly changes over the cases in the fixed effects estimates (columns 2-4 of Table 6.8) while the magnitude of the change between the cases is much stronger in the random effects estimates. The coefficient that attached to the Theil's index is largely declined through the estimation cases in the random effects estimate (from 0.9195 in the case of no controls and dummies to 0.3041 in the case of with controls but no dummies and to only -1.3994 in the case with both controls and dummies as reported in raw 3 of columns 5, 6 and 7 of Table 6.8). To investigate the correlation between growth rate and initial level of per capita income, the third row of Table 6.8 indicates that in the fixed effects estimates, there is a negative and not a statistically significant relationship between economic growth rate and log level of GDP per capita for the cases of with the controls but no time dummies and with both the controls and time dummies (columns 3 and 4 of Table 6.8), and positive correlation between the rate of growth and log level of GDP per capita (column 2 of Table 6.8). Meanwhile, in the random effects estimates we found that the relationship between the log level of GDP per capita and growth rate is positive in the cases of no controls and with controls but no dummies, and negative in the case with both controls and dummies. However, the relationship is not significant except for the case with controls but no regional and time dummies where the coefficient attached to Theil's index of inequality is statistically significant at 7.5% level. The low value of R^2 provided in the last row of Table 6.8 may indicate that the model did not explain well the relationship between growth rate and regional income inequality, except for the random effects estimate for the case with both the controls and dummies (R^2=26%).

We reported in Table 6.8 the estimates by both the fixed effects and random effects methods. Now, we set to test the validity of the assumptions underlying each method in order to find out which is an appropriate method of estimation for a proper conclusion. We use Hausman specification test comparing fixed-effects estimates reported in columns two, three and four versus random-effects estimates reported in columns five, six and seven of Table 6.8 for three cases: without both controls and time dummies, with controls but no regional and time dummies and with both controls and regional & time dummies respectively. We test the null hypothesis that the random effect is preferred against the alternative that random effect is not consistent (in other statement, under the null hypothesis, the random effects estimator is consistent and efficient against the alternative that it is inconsistent and inefficient). The test results are reported in Table 6.9. Based on the test results, for the case of random effects estimates without controls and regional and times dummies, and the case of random effects estimates with controls but no regional and time dummies, the calculated Chi square is larger than the critical value (columns two and three of Table 6.9). Thus, we are able to reject the null at least at 5% level of significance. Thus for these cases, random effect estimators are inconsistent and inefficient, and the fixed effects method is preferred. For the remaining case: random effects estimates with both controls and regional and time dummies, we are unable to reject the null because calculated Chi square is well below critical Chi square (column four of Table 6.9). We thus conclude that for this case, random effects technique yields consistent and efficient estimators. This test indicates that random effects technique for the model with the inclusion of controls variable, regional and time dummies is and best explained the model and most preferred even though the values of AIC and BIC in the models with both controls and dummies are slightly higher than the models without both controls and dummies, and with controls but no dummies (see Table 6.8).

Table 6.8. Basic fixed effects and random effects regressions: entire growth period

	Basic Fixed Effects			Basic Random Effects		
	No Controls	Controls	Controls and time & regional dummies	No Controls	Controls	Controls and time & regional dummies
Logy	0.0012	-0.1128	-0.0795	0.0871	0.2566*	-0.0331
	(0.126)	(0.201)	(0.221)	(0.120)	(0.144)	(0.163)
Inequality	3.5025	3.4043	3.6835	0.9195	0.3041	-1.3994
	(2.6316)	(2.669)	(2.776)	(1.615)	(1.212)	(0.862)
Invest		0.003	0.003		0.0003	0.0008
		(0.003)	(0.003)		(0.003)	(0.0025)
Second		0.0067	0.0095		-0.0115	0.0016
		(0.011)	(0.013)		(0.01)	(0.0096)
Old		0.097	0.1025		-0.388**	0.0546
		(0.739)	(0.739)		(0.171)	(0.126)
Regional dummy			-			0.6875
						(0.066)
Time dummy			-0.1249			0.1463
			(0.338)			(0.273)
Number of observations	1031	1031	1031	1031	1031	1031
Adjusted R^2	0.02	0.06	0.07	0.05	0.07	0.26
AIC				5012.38	5017.12	5018.98
BIC				5027.20	5046.75	5053.54

Notes: Standard errors in parentheses. ** indicates statistically significant at 2.5 % level; * statistically significant at 7.5% level. Logy is the log per capita GDP; Inequality indicates the regional Theil's index; invest stands for ratio of public investment in GDP; second is the share of people with secondary education (this variable indicates human capital available in region); old stands for share of elderly. Dependent variable is log of annual growth rate. Estimation is by fixed-effects and random-effect techniques.

Table 6.9. Hausman specification test of random effects method versus fixed effects method

	No controls and time dummies	With controls but no time dummies	Without controls and time dummies
Degrees of freedom	2	5	6
Calculated χ^2	5.77	20.54	4.79
Critical χ^2	4.61	16.75	14.45
P-value > Chi Square	0.056	0.001	0.571

Based on the estimation results using pooled OLS and basic random effects techniques presented in respective Table 6.7 and Table 6.8, we find that the both techniques yield a same level of R^2 (of 0.26). However, if we compare the values of AIC and BIC[54] between the two estimation techniques: pooled OLS and ransom effects (the last two rows of Table 6.7 and Table 6.8), we find that AIC and BIC values produced by the pooled OLS are much smaller than the those produced by the random-effects technique. Thus, we conclude that the model with pooled OLS technique of estimation is preferred.

6.8. Conclusion and discussion

In this chapter, we have measured and analyzed the regional income inequality, and explored the evolution of per capita income disparities and its decomposed components

[54] Given any two estimated models, the one with lower value of AIC and BIC is preferred.

and its relationship with economic growth among regions in Vietnam in the period from 1990 to 2006.

Firstly, we have measured and analyzed the regional income inequality in Vietnam. Our findings have confirmed that generally the regional income inequality in the country moderately increases over the time. The results of decomposition of total income inequality indicate that, over time with the process of economic development, the distribution of income within region becomes more and more equal compared with the initial period while the distribution of income between regions becomes more unequal. However, the size of the within region inequality is different among regions in the country.

To explore the relationship between inequality and growth, we have used several different techniques based on and adapted from the studies on the issue by Barro (2000), Pinizza (2002) and Forbes (2000). The techniques used to estimate the growth model are panel data OLS technique, pooled OLS technique and the fixed-effects and random effects methods. Our findings indicate that there are different outcomes in the relationship between regional economic growth and regional inequality based on the method of estimation.

The estimates from the panel data regression for both the entire growth period and for different growth episodes have found a positive relationship between the rate of economic growth and the initial level of per capita income in all the cases: without controls and regional dummies, with control variables and no regional dummies, and without both the controls and regional dummies. However, the relationship is statistically significant for only the case without both control variables and dummies in the growth model for the entire growth period estimation, and for the case of estimate for five-year growth episode starting from 2001 and for ten-year growth episode starting from 1996.

In the pooled OLS estimates for the growth model, we find that there are negative relationships between growth and inequality for all the cases: without both controls and regional dummies, with the controls and no regional dummies, and with both controls and regional dummies (Table 6.6). However, the relationship is found to be significant only for the case with both controls and regional dummies (row three, column three of Table 6.6).

For the fixed effect and random-effects estimation, the test for the validity of the method used in this study conducted above indicates that the fixed effects method is appropriate for the cases of without both the controls and dummies and with controls but no dummies in the growth model while the random effects technique is more appropriate and more preferred for the case with both the controls, regional and time dummies. The fixed effects estimates for all the cases given in columns two, three and four of Table 6.8 and random effects estimates for the case of without both the controls and dummies and with controls but no dummies given in columns five and six of Table 6.8 confirm that there is a positive but not statistically significant relationship between growth rate and regional inequality. The random effect estimates for the case of with both controls and dummies (Column seven of Table 6.8) indicate that there is a negative and no statistically significant relationship between growth and inequality in Vietnam. Interestingly the estimates for different cases by both the fixed effects and random effects methods reveal that there are both the positive relationship (for most cases) and negative relationship (only for the case of with both the controls and regional dummies in random effects estimate) between growth rate and initial level of per capita income (see raw 3 of Table 8.8 above). The positive and statistically not significant relationship is found when the no controls and dummies are included in the growth model using fixed effects technique and when the no dummies are included in the model using random effects method while the negative and not statistically significant relationship is found in the remaining cases. The findings of the positive and negative relationship between inequality and growth by different methods of estimation are also in line with the findings from the analysis of regional inequality and regional growth in both the developed and developing countries.

The empirical results of the relationship between regional inequality and regional growth show the fact that the estimated relationship between inequality and growth changes when using different estimation techniques. However the estimation using panel data regression, the fixed effects and most cases of random effects methods seem to support theoretical argument and empirical findings by , among others Alesina and Rodrik (1994), Perssion and Tabellini (1994) and Parotti (1996) that there is a positive correlation between growth and inequality for a developing economy like the Vietnamese, while pooled OLS regression yield the results of the negative relationship between the two variables equality and growth. The same results are found in many empirical studies by, among others, Forbes (2000) and Li and Zou (1998).

Moreover, we do not think that the results of this chapter are strong enough to justify that the study finds empirical support for some specific channel linking inequality to growth because the different results are obtained by using different techniques of estimation and the length of the growth period is not really sufficient long for the empirical testing. However, based on the results of Hausman test, we prefer the random effects model with both controls and dummies with the value of Adjusted R^2 is much higher than its counterpart by the fixed-effect model (0.26 versus 0.07) (see Table 6.8). Based on the estimation results using pooled OLS and basic random effects techniques presented in respective Table 6.7 and Table 6.8, we find that the both techniques yield a same level of R^2 (of 0.26). However, if we compare the values of AIC and BIC between the two estimation techniques: pooled OLS and ransom effects (the last two rows of Table 6.7 and Table 6.8), we find that AIC and BIC values produced by the pooled OLS are much smaller than the those produced by the random-effects technique. Thus, we conclude that the model with pooled OLS technique of estimation is preferred.

Chapter seven
Regional Income Divergence and the Effects of Spatial Integration

7.1. Introduction

In the previous chapters we have examined multidimensional regional welfare inequality for two components of welfare namely education and health, and analyzed regional income inequality. In these chapters we have also carried out the tests for spatial dependence among dimensions of the health component of welfare, explored a possible causal relationship between the two dimensions of education, and analyzed the correlation between regional income inequality and economic growth. To further explore the economic growth tendency, in chapter seven we go on to examine the patterns of regional income growth in Vietnam using provincial - level income and growth data. As findings from chapter six reveal that regional income inequality have been increasing across regions in Vietnam since the country has opened its economy to greater regional and international economic integration in 1986. However, the pattern of increasing regional economic growth and regional income inequality has not been fully and systematically explored and analyzed. Thus, in this chapter we will scrutinize whether regional income diverges or converges across the provinces of the country, and whether the growth pattern of the GDP per capita in Vietnam is in line with the findings in the literature of the development economics. As many studies on the issue for developing and transitional countries like Vietnam , China and Russia found evidence for both convergence and divergence patterns in regional income in the initial period of economic growth as the evidence for divergence is found by studies of China's and Russia's economy by, among others, Pedroni and Yao (2005), Zhang *et al* (2001), and Bradshaw and Vartapetov (2003), while the evidence for convergence is found by the studies of China's and Russia's economy by Chen and Fleisher (1996), Cei *et al*. (2002), Demurger *et al.* (2002), and Weeks and Yao (2003). The regional divergence or convergence in incomes may be due largely to geographic factors, or to differences to which provinces promote open-door policies for foreign investment projects and foreign trade to promote economic growth. Thus, in this study we will focus on more fundamental causes linked to convergence/divergence in regional incomes caused by geographic factors that are considered as a connection to the mechanism for rapid growth. We do not consider the case related to differences to which provinces promote open-door policy (the policies to attract foreign direct investment and to open the economy to foreign trade in order to promote economic growth) because of a lack of information on the pro-growth and open-door policies of provinces.

The picture of Vietnam's economic growth has changed significantly following the *doi moi* initiated in 1986, which has introduced economic incentives in all sectors of the economy and opened the economy to foreign trade and investment. Before the *doi moi*, real per capita income grew at around 2.2% per annum in average for the time span of almost 10 years, after the unification of the country in 1975 to time of the *doi moi* embarkation. The economic growth during this period was quite erratic. The economy was in recession in the first half of the 1980s with failures in agriculture and industry caused by mismanagement of the economy. *doi moi* started in 1986 and has brought about drastic changes in the growth of economy with large influx of foreign direct investment flow and an increase in bilateral trade between Vietnam and other countries all over the world. With these positive changes, according to the statistics archived in the GSO office of Vietnam

(GSO, various years) the Vietnamese economy has unleashed a period of unprecedented rapid and steady growth, at more than 6% in annual average in real terms during the period between 1990 and 2006. The contrast in the growth experience between the pre- and post-*doi moi* periods can be seen in Figure 7.1, which depicts real log per capita income in average for the 61 provinces going back to 1975. Note that the data of the real per capita GDP growth for the period from 1975 to 1990 were at the national average. The data was not available for each individual province during this period. As the picture shows, the periods of the pre and the beginning year of the *doi moi* are characterized by the unsteady growth with relatively wide fluctuations in per capita income. Since the *doi moi*, particularly since 1990, the growth rate of per capita GDP showed a steady state as the data and the figure indicate.

For a long time, the study of convergence process at the regional level (between regions within a country) and international level (between countries) has drawn a lot of attention from researchers worldwide. Regions extensively studied are of the US (Rey and Montoury, 1999, Miller and Genk, 2005), the European regions (EU-15) (Lope'z-Bazo et al., 1999, Le Gallo *et. al*., 2003, Arbia and Piras, 2004), income convergence/divergence among provinces/states in developing and transitional countries like Vietnam such as China (Zhang *et al*., 2001, and Pedroni and Yao, 2005) and Russia (Solanko, 2003; Bradshow and Vartapetov, 2003), and income convergence/divergence among provinces/states in a developing and middle income countries likes Brazil (Azzoni, 2000). The study on convergence among European countries including old members of EU and the new members from Eastern Europe has shown two features. First, it is widely accepted that the rate of convergence among the regions is quite low[55]. The second feature refers to the geographical distribution of economic activities. Several studies (by, among others, Lope'z-Baro *et. al*., 1999; Le Gallo and Ertur, 2003; Magrini, 2004) have observed that European regions were divided into two groups: rich in the North and poor in the South. This observation can be linked to several results of the New Economic Geography, that there exists agglomeration and cumulative processes which spatially determine the location of economic activities (Krugman, 1991). For example, it is thought that a region which bordered with rich regions is more likely to have a higher economic growth than regions that are surrounded by poor ones. This point referred to spatial interactions among neighboring regions, which is especially relevant at the regional level.

Despite its relative importance to the regional studies, however, this issue was completely ignored by most researchers working on this topic so far, because of the fact that capturing the spatial effects in regression usually generates a lot of problems for analysis. However, with the spatial econometrics tools developed by Anselin (1988) and other econometricians, we already have a solution for this difficulty. The spatial econometric tools allow one to capture the spatial dimension in an econometric regression. Problems caused by taking space into consideration in the econometric models are of two dimensions: spatial autocorrelation and spatial heterogeneity. In other words, space interferes in two different ways on regression's estimators. These are called spatial effects. When we work on provincial regions in Vietnam, we of course think of correlation among observation units. In other words, the value of a variable in a location is partly affected by the value of the same variable in adjacent locations. This effect is called spatial

[55] This rate of convergence (<1% per year) is considered low compared with the rate generally agreed in the literatures as 2% per year (see, among others, Barro and Sala-I-Martin, 1991, 1995; Armstrong, 1995; Ertur et al., 2006).

autocorrelation and it refers to the presence of dependence among regions. In addition, if we think of the fact that economic relationship and behavior differ across regions as, for instance, rural versus urban behavior and performance, we would talk of the second spatial effect called spatial heterogeneity. In Vietnam, there is still a lack of studies on this topic. This chapter therefore is very pioneering in looking in this matter in exploring regional income divergence in Vietnam and the role of space in its course.

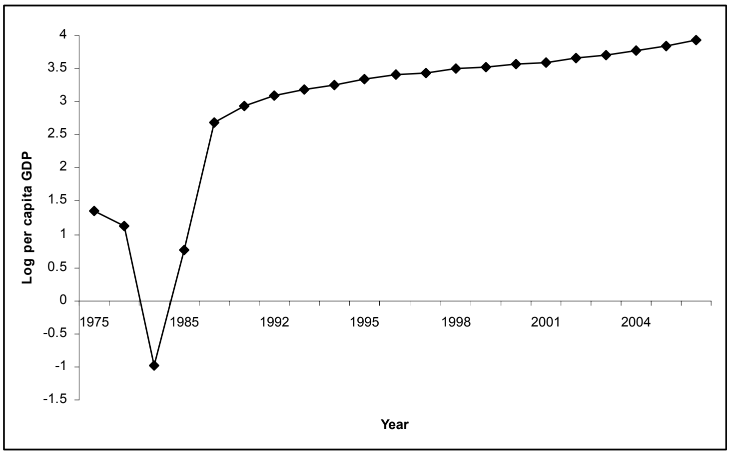


Figure 7.1. Average provincial per capita GDP in US $ at constant 1994 prices

7.2. Informal data analysis of income convergence and divergence in Vietnam

This section examines the raw data informally. The data used consist of the provincial GDP for sixty-one provinces of Vietnam for the seventeen-year period, going from 1990 to 2006. The provincial GDP that we use is provided by the General Statistical Office of Vietnam based on the yearly basic statistics. In order to study income convergence, we use total population in combination with GDP and deflated by provincial GDP deflators to generate real per capita GDP by province. The GDP deflator takes the price of 1994 as 100.

The most important achievement of the Vietnam economic reform (*doi moi*) is its outstanding economic growth performance. Over the last 15 years Vietnam's economy has been growing at an impressive growth rate, enabling per capita GDP to almost quintuple, from USD 151 in 1984 to USD 722 in 2006 (figures are measured at the 1994 prices). At present, Vietnam is one of the fastest-growing economies of the world. Table 7.1 below displays the average annual growth rate of the real per capita GDP of all the provinces during the 1990-2006 period which indicates that the growth rate attained is high and steady over the last 15 years. For the country as a whole, the average growth rate of per capita GDP over the period was 7.8% while the average rate of per capita GDP growth before the *doi moi* was very low, somewhere at 3.5% annually during the 1975-90 period (General Statistical Office (GSO), several years). Because of a lack of statistics on economic performance for the period before the *doi moi*, we are not able to provide the illustrations of the per capita GDP growth rates by provinces herewith. However, as the indicated and elaborated in chapter two, the economy of Vietnam was in a situation of deep recession during the time before the start of the *doi moi*.

The acceleration of economic growth has affected all provinces of Vietnam. All provinces average annual growth rates during the *doi moi* are very high, at 7.8%, much higher than the rates at the time before the *doi moi* (1976-1990 period), which stood at around 2.2%. In particular, most metropolitan cities and provinces that attracted many foreign direct investment (FDI) projects such as Hanoi, Hai Phong, Vinh Phuc, Bac Ninh, Da Nang, Binh Duong, and Ho Chi Minh City, have sustained quite high rates of per capita GDP growth compared with the rates recorded by other remaining provinces (see Table 7.1). However, the annual growth rates of the per capita GDP over time during the *doi moi* period vary significantly across provinces. As reported in Table 7.1, the standard deviation of the provincial growth rates between the province with the lowest rate and the one with the highest rate is more than 300%. At annual average, the slowest growing province, Dak Lak, only grew at 4.02% while the fastest growing province, Vinh Phuc, grew at 13.93% during the period from 1990 to 2006. Most provinces recorded high rates of growth that lie between 5% and 10% but there were six provinces that have managed to grow at double-digit rates[56]. Obviously, five of these six provinces were located either in the northern part or in the southern part of the country which were considered as richer regions in Vietnam. Only one of them, Da Nang, is in the South Central Coast, which is considered as the economic centre of the central coastal region of the country. Interestingly, however, in geographical terms, it is not the case that all provinces in the North or in the South have been growing faster than the provinces in the remaining regions of the country. For example, looking at the statistics provided in Table 7.1, the rates of growth average from around 7% to 9% in most provinces in all the regions, not only in the North, in the South and in the Coastal lines, but also in the Central Highlands as well. Thus, it seems that, at least in this study, there is not any evidence of geographical factors found to have effects on regional economic growth. In other words, spatial interactions among factors in neighboring regions may not appear to be the necessary condition for enhancing economic growth.

We now illustrate by a graphic analysis some of the key features of the provincial growth process during the *doi moi* period based on the available data. These illustrations help us gaining an overview of the tendencies towards income convergence or divergence among 61 provinces of Vietnam since the start of the *doi moi* before carrying out more formal tests for convergence using econometric tools.

[56] These provinces are Hanoi, Vinh Phuc, Ha Giang, Da Nang, Binh Duong and Ba Ria-Vung Tau, with the average annual growth rates of per capita GDP over the 1990-2006 period were respective 10.31%, 13.93%, 10.68%, 10.35%, 10.68%, and 10.41%.

Table 7.1. Growth performance during doi moi by provinces in Vietnam

Average annual growth rate of per capita GDP during *doi moi* period (1990-2006)			
Province	Rates	Province	Rates
Hanoi	10.31	Da Nang	10.35
Hai Phong	7.24	Quang Nam	9.16
Vinh Phuc	13.93	Quang Ngai	7.98
Ha Tay	7.53	Binh Dinh	7.37
Bac Ninh	9.79	Phu Yen	9.24
Hai Duong	7.61	Khanh Hoa	8.96
Hung Yen	8.87	Kon Tum	8.69
Ha Nam	7.03	Gia Lai	6.45
Nam Dinh	6.04	Dak Lak	4.02
Thai Binh	6.27	Lam Dong	6.69
Ninh Binh	7.69	Ho Chi Minh City	7.87
Ha Giang	10.68	Ninh Thuan	5.13
Cao Bang	9.50	Binh Phuoc	6.24
Lao Cai	5.98	Tay Ninh	9.45
Bac Kan	7.20	Binh Duong	10.68
Lang Son	7.84	Dong Nai	9.26
Tuyen Quang	7.10	Binh Thuan	8.39
Yen Bai	7.95	B. Ria - V. Tau	10.41
Thai Nguyen	6.56	Long An	6.42
Phu Tho	6.83	Dong Thap	6.55
Bac Giang	6.20	An Giang	8.05
Quang Ninh	8.82	Tien Giang	7.04
Lai Chau	5.70	Vinh Long	7.25
Son La	7.27	Ben Tre	7.76
Hoa Binh	6.72	Kien Giang	6.92
Thanh Hoa	6.10	Can Tho	7.49
Nghe An	6.70	Tra Vinh	7.76
Ha Tinh	5.99	Soc Trang	7.02
Quang Binh	7.53	Bac Lieu	8.75
Quang Tri	7.43	Ca Mau	6.45
Thua Thien - Hue	7.59		

Source: GSO, Vietnam (Statistical Yearbook)

To illustrate the provincial growth pattern we examine Ba Ria-Vung Tau (the richest province in term of GDP and GDP per capita that is famous for its rich natural resources as crude oil and the home for many crude exploitation companies) in comparison with those province which are catching up to Ba Ria-Vung Tau in term of per capita GDP such as Hanoi, Ho Chi Minh city, Binh Duong and Da Nang, etc., and those that are falling behind Ba Ria – Vung Tau in the level of per capita GDP such as provinces in the Northwest and Central Coast. Among the rich provinces that are considered catching up to Ba Ria-Vung Tau to be illustrated herewith are Hanoi, Ho Chi Minh City, Hai Phong and Binh Duong. Figure 7.2 shows the extent to which the average per capita income of these provinces is approaching the one of Ba Ria-Vung Tau. Figure 7.2 shows that provincial average incomes of all provinces in this group excluding Ba Ria-Vung Tau seem to converge and approach each other while that of Ba Ria-Vung Tau lightly diverges. Figure

7.3 illustrates average GDP per capita of the Ba Ria Vung Tau in perspective with the mean GDP per capita of other two clusters such as (i) the cluster of richest group without Ba Ria – Vung Tau that comprises of rich provinces in terms of GDP per capita and includes Hanoi, Hai Phong, Ho Chi Minh City and Binh Duong, and (ii) the clusters of provinces having a GDP per capita that are falling behind the per capita GDP level of the richest. We call them the falling-behind groups. The falling-behind groups include the cluster of provinces of the North Central Coast, the cluster of provinces of the North East, excluding relatively higher income provinces as Thai Nguyen, Phu Tho and Quang Ninh, and the cluster of provinces of the entire North Western region. We observe from Figure 7.3 that the per capita incomes seem to be diverging away among the clusters considered. The richest province: Ba Ria-Vung Tau and the high-income group which consists of Hanoi, Ho Chi Minh city, Hai Phong, Da Nang and Binh Duong are diverging away from the falling-behind groups such as the North Central Coast, the North East and the North West. There as only a minor sign of converging within the low-income group as lowest three lines in Figure 7.3 represent growth patterns for three regions: the North Central Coast, the Northeast and the Northwest are observed to be closer to each other in the end than at the start.

Many studies by the World Bank (various years) of the Vietnamese economy tend to categorize provinces simply into mountainous and interior provinces versus low-land and coastal provinces, since historically there has been a significant difference in their relative growth rates. However, this trend seems to be directed toward another direction as presently provincial growth rate is different between provinces with many foreign direct investment (FDI) projects and those without or less FDI projects. In fact, FDI projects usually accrued into the lowland and coastal provinces due to its advantages in labor force, in transportation and infrastructure. Thus, the higher rate of growth still seems to happen in the lowland provinces throughout Vietnam except for several cases that have advantages in industrial crop production, mining exploitation or tourist attractive sites such as provinces in the Central Highlands and some provinces in the mountainous Northern part of the country.

If we refer to convergence as the narrowing of inter-provincial income differences, it seems that there is a relatively little overall convergence among clusters of all provinces when we examined the income pattern over the entire sample period. Furthermore, as we have seen in an illustration in Figure 7.3 that regional clusters as well as rich or poor provinces fail to converge to each other. It means that interregional income disparities are likely to continue to diverge.

In short, the informal data analysis of the economic growth based on the provincial growth rates in Vietnam has provided the reader a preliminary look at the scenery of the growth pattern and convergence or divergence in per capita GDP among provinces of Vietnam. The analysis has revealed the mixed trends in the patterns of per capita GDP growth in the country as evidence on convergence among the high income provinces (Figure 7.2) and among the poor provinces (Figure 7.3) was found, and evidence on divergence was found between Ba Ria-Vung Tau and other high income provinces (Figure 7.2) and among Ba Ria - Vung Tau, the cluster of high income provinces and remaining clusters of poor provinces in the North Central Coast, in the North East and in the North West. These findings again proved that the growth patterns in the developing and transitional economies like Vietnam and China has experienced different path regarding which clusters we consider and the way we choose to make groupings for analysis.

The informal analysis based on the growth rate presented in Table 7.1 above and on the graphic presentation as in Figure 7.2 and 7.3 below is useful in informally developing a sense of the extent of the convergence or divergence present in the data used for the analysis. In order to find the answer to the question whether the data is in fact consistent with long run convergence or divergence in a formal sense, we turn to the next section in which a more systematic empirical analysis and tests are implemented.

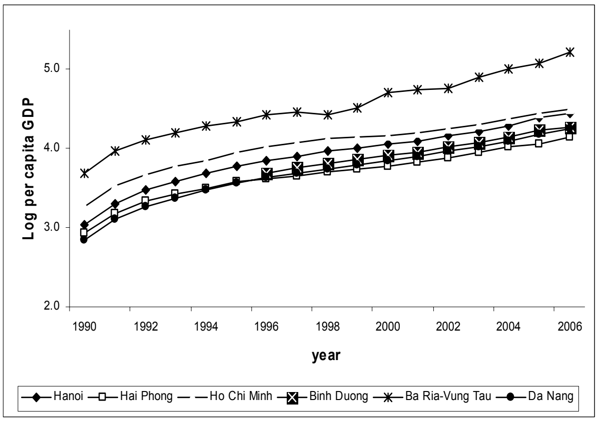


Figure 7.2. Average provincial per capita incomes of the high income provinces relative to richest province Ba Ria-Vung Tau

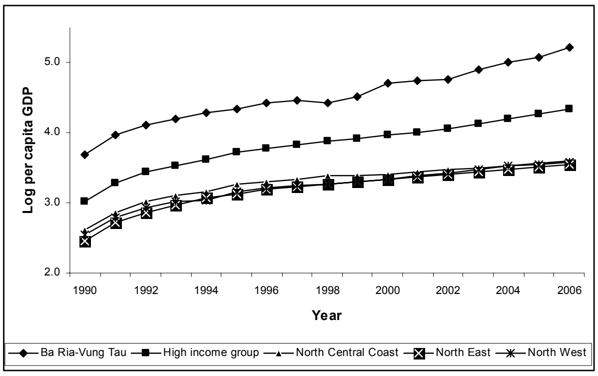


Figure 7.3. Richest province and high income group versus poor and falling-behind clusters

7.3. General formulation of the panel based tests for convergence

In section 7.3.1 we first elaborate various notions of income convergence and methods of modeling. Various tests for convergence are given in section 7.3.2. In section 7.3.3 we continue with a discussion of the role of spatial effects in the analysis of regional income convergence using econometric tools.

7.3.1. The concept of convergence and modeling convergence

To more rightly investigate whether the data are consistent with convergence or divergence, we begin with a formal definition of what we mean by the concept of convergence as there are two concepts of convergence regularly mentioned in the literature of economic growth: the beta-convergence and the sigma-convergence. Then, we develop several frameworks to test the presence of convergence process among regions in Vietnam based on the work by Barro and Sala-I-Martin (1995), and Evans (1998).

We start with the definition of a convergence process. A convergence process is defined as a negative correlation between the initial income and the average rate of economic growth with or without isolating the effects of other factors of economic growth (Barro and Sala-I-Martin, 1995). Then, we proceed to elaborate two kinds of convergence: beta-convergence and sigma-convergence that are commonly used in the literature of development economics. The first form of convergence, which has primarily been the focus of macroeconomics, occurs when poor regions (or countries) grow faster than rich ones, resulting in the former eventually catching up to the latter in per capita income levels. Such a situation is always referred to as beta-convergence. The second convergence concept refers to the decline in the cross-sectional dispersion of per capita incomes. In the literature of economic growth, several different measures have been used to examine this form of convergence including the un-weighted standard deviation and coefficient of variation of the log of per capita income (see Carlino and Mills, 1996b; Bernard and Jones, 1996, among others). This form of convergence is referred to as sigma-convergence which has attracted much attention from the researchers (among others, Kuznets, 1955; Easterlin, 1960a, 1960b; Williamson, 1965; Amos, 1988). In other words, sigma-convergence is the process when country-wise inequality tends to reduce in the course of time. Generally, convergence of the first type tends to generate the second type of convergence: when the poor grow faster than the rich ones, the result is the reduction in the differences in per capita income across individuals regions.

Vietnam is in its initial stage of development. According to the development economic literature (Kuznets, 1955; Barro and Sala-I-Martin, 1995), the regional GDP per capita in Vietnam is still diverging as the poor regions are still lagging behind the rich regions in terms of growth rate. Therefore, to be more precisely, the gross regional product (GRP) per capita across regions in the country is in divergent process. For this reason, we will not look at the sigma-convergence because inequality is still increasing in the country. For the beta-convergence, which is defined as a process that poor regions grow faster than the rich ones resulting in former eventually catching up the latter in per capita income levels. For this type of convergence, there are two distinct convergence notions under this form of convergence to be considered: absolute (unconditional) β-convergence and conditional β-convergence[57]. The concept of β-convergence directly comes from the neoclassical economic theory. This theory stipulates that in the long run per capita GDP growth depends only on exogenous technical progress. When we apply this idea to several economies, there are different situations that must be distinguished. If the economies share the same structural characteristics (in terms of human capital, saving rate, production

[57] The literature makes a distinction between conditional and unconditional (absolute) convergence, with the former pertaining to the partial correlation of the growth rate and initial level after controlling for a number of additional variables reflecting differences in equilibrium wages and technologies (Barro and Sala-I-Martin, 1992).

function, etc), the convergence in GDP per capita and in growth rate is present (Debarsy and Eutur, 2006). However, if structural characteristics differ among economies, only a convergence in growth rate is observed. In this case we talk of absolute (unconditional) β-convergence because the long-run equilibrium is the same for all economies. The concept used is thus conditional β-convergence. To test absolute (or unconditional) convergence process we use the following specification developed by Barro and Sala-I-Martin (1995) for cross-sectional data:

$$y_T = \alpha + \beta y_0 + \varepsilon \qquad (7.1)$$

Where:

y_T is the $(N,1)$ vector of log average annual growth rate of per capita GDP,
N is number of observations,
T (subscript of dependent variable y) is the number of periods (years) or sample size under study, and
y_o is the log per capita GDP in the initial period (year).

To test absolute convergence process, we assume that ε is a vector of normally distributed random error terms with means 0 and homoskedastic variance σ^2. Convergence is observed if β is negative and significant. In fact, the growth rate is then negatively correlated with the level of per capita GDP.

To test for conditional β-convergence hypothesis, we need to include conditioning variables. The equation for the test is as follows:

$$y_T = \alpha + \beta y_0 + \gamma S + \varepsilon \qquad (7.2)$$

with all the variables and notions are defined as above and S is a matrix of log conditioning variables. This matrix consists of state variables such as the human capital stock (employment rate), public spending on GDP, fertility rate and education enrolments (human capital). Conditional convergence is present if β is negative and significant when S is kept constant.

Based on equations (7.1) and (7.2), two more parameters can be computed. The first refers to the speed of convergence and the second refers to the time necessary to reach the steady state, known in the literature as the *half-life*. The equations to compute the speed of convergence and the *half-life* are as follows:

$$\xi = -\log(1 + T\beta) / T \qquad (7.3)$$

and

$$\tau = -\log(2) / \log(1 + \beta) \qquad (7.4)$$

Where:

ζ is the speed of convergence,
τ is the half-life,
T is the number of periods (years) or sample size under the study (as defined in 7.1), and
β is the convergence coefficient estimated from the equations (7.1) and (7.2).

However, the conventional *β-convergence* treats regions as 'isolated islands' (Quah, 1996). It does not capture the fact that the economic performance of a region depends upon those of other regions. In fact, the development of each region is closely correlated with the development of, at least, neighboring regions. In this context of spatial dependence, the distribution of regional per capita income is unlikely to be spatially independent. When models are estimated for cross-sectional data, the lack of independence across these units may cause critical problems of model misspecification and biased estimators when ignored (Anselin, 1988). In addition, as discussed in the section of explanatory analysis of the data, we use panel spatial data for this study. One of advantages of the panel spatial data is that we can use spatial econometric methods because each observation is geographically located. In fact, until recently, in many studies, countries and regions were treated as independent economies without any interaction. Quah (1996) has raised concerns about the neglected spatial dimension of the *β*-convergence. However, at the regional level, spatial spillovers are of interest given that each region is likely to interact with neighboring regions. We therefore will take space into consideration. To do this, the explanation of how to take spatial effects into an econometric regression is considered in section 7.3.3.

7.3.2. Convergence testing

To test for convergence, we use several traditional tests used by many studies (see, among others, Anselin, 1998; Debarsy and Urtur, 2006; Barro R. J. and X. Sala-I-Martin, 1992; and Arbia and Piras, 2004) such as Moran's I statistics and Lagrange Multiplier (LM) test. In addition to these tests, another test for convergence using unit root test is also discussed and applied in this chapter.

- ***Moran's I test***

Moran's test plays a predominant role in applied empirical econometric research. Traditionally, this test is used to test for spatial dependence. However, it is interesting that the measure of spatial dependence also tends to commove with the measure of income dispersion. The spatial association of the income level and the growth of per capita income is analyzed by Moran's correlation coefficient. So, in many studies, (see, among other, Anselin; 1988; Barro R. J. and X. Sala-I-Martin, 1992), Moran's statistics is used to test for the growth convergence in income level. Moran coefficient is given by:

$$I = \frac{e^{'}.W.e / S_w}{e^{'}.e / N} \qquad (7.5a)$$

with e as a vector of OLS residuals, N is number of regions, W is the standardized spatial weight matrix, and S_w is the sum of all elements in the spatial weights matrix *W*. For a row-standardized *W, S* simplifies to *N*, so that *N/S=1* and can be ignored. Therefore (7.5a) becomes[58]:

$$MI = e'We / e'e \qquad (7.5b)$$

For normal error terms the distribution of the standardized Moran I statistic is shown to be asymptotically distributed as χ^2 with one degree of freedom. The statistic shows a striking similarity to the familiar Durbin-Watson test. In the spatial case, the critical value will depend not only on *X*, but also on the asymmetric matrix of *W*, which is

[58] See note number 39 in chapter four.

clearly not very practical (Anselin, 1991). Also, Cliff and Ord (1972, 1973, 1981) show that the asymptotic distribution of Moran *I* based on least-square residuals corresponds to a standard normal distribution.

- ***Lagrange Multiplier (LM) based test***

When the spatial regression models are estimated by maximum likelihood (ML), inference on the spatial autoregressive coefficients may be based on Wald or asymptotic t-test (from the asymptotic variance matrix) or on a likelihood ratio test (see Anselin 1988, chapter 6; Anselin and Bera 1998). Both approaches require that the alternative model (i.e., the spatial model) be estimated. In contrary, a series of test statistics based on the Lagrange Multiplier (LM) principle require only estimation of the model under the alternative. LM test also allow for the distinction between a spatial error and a spatial lag alternative.

In addition, In Monte Carlo experiments Anselin and Florax (1995) have also shown that the Bera-Yoon LM tests have high power in discriminating between spatial lag and error dependence. The basic idea of these tests consists in correcting the LM test statistics for spatial error and lag dependence. According to Anselin and Florax (1995: 25-29), LM(err) and LM(lag) can be presented as:

$$LM_{(err)} = (e^{'}.W.e/\hat{\sigma}^2)^2 / tr(W^{'}W + W^2) \qquad (7.6)$$

and

$$LM_{(lag)} = (e^{'}.W.y_T/\hat{\sigma}^2)^2 / [(Wy_0\hat{B})^{'} MWy_0\hat{B}/\hat{\sigma}^2 + tr(W^{'}W + W^2)] \qquad (7.7)$$

with the ML estimator $\hat{\sigma}^2 = e^{'}.e/n$ for the error variance σ^2 and $M = I - y_0(y_0^{'}y_0)^{-1}y_0^{'}$ known from regression analysis as a projection matrix. To distinguish among alternatives, the spatial lag dependence is eliminated from LM(err) by extracting a function of $e^{'}.W.y_T$, whereas a function of $e^{'}.W.e$ is subtracted from LM(lag) in order to control spatial error dependence (see, among others, Bera and Yoon, 1993; Anselin and Florax, 1995). This means that the adjusted LM error test (LM(err) robust test) responds to spatial error dependence but not to spatial lag dependence. In contrary, the adjusted LM lag test (LM(lag) robust test), is expected to indicate spatial lad dependence but not spatial error dependence. Under the assumption that the errors are normally distributed, both the LM(err) and LM(lag) as well as their robust counterparts tests follow a χ^2 distribution with one degree of freedom.

- ***ADF (IPS[59]) test for unit root***

Next, another notion of convergence test discussed by Evans (1998) will be employed in order to consolidate the above tests using the framework developed and used by Barro and Sala-I-Martin (1995). This approach asked whether or not the long-run forecasts for output differences converge as the forecasting horizon increases, which implies that the long-run income gap between any two provinces must be stationary. To formalize this idea empirically, we can characterize the convergence as follows. If we

[59] The ADF (IPS) test for unit root in the analysis of income convergence used in this thesis is developed based on the approach used by Im, Pesaran, and Shin (2003). IPS stands for the combined initial letter of the names of these three researchers.

denote y_{it}, the logarithm of per capita output for province i at time t, as difference stationary, and thus exhibits unit root behavior individually. Then any pair of provinces i and j are said to converge pair wise if the difference $(y_{it} - y_{jt})$ is stationary so that y_{it} and y_{jt} are co-integrated. Analogously, convergence among members of a larger group of provinces is then defined by requiring that every pair within the set exhibits convergence. One obvious advantage to this method of focusing on the properties of long-run income gaps as the criteria for convergence is that it directly allows us to infer long run forecasts for the absence or presence of income disparities between provinces.

For the empirical study of the case of Vietnam's economy, for the group of provinces, we might test this condition pair wise for all provinces within the sample and then require that the condition holds for each possible pair of provinces. A disadvantage of the approach is that conventional tests for co-integration tend to have low power for such a relatively short sample, and so the probability of failing to reject the null hypothesis of no co-integration for at least some pairs would be quite high regardless of the true relationship (Evans, 1998). Fortunately, Evans (1998) shows that it is possible to translate these criteria into a single criterion that should apply to the group as a whole when interpreted as a panel. Particularly, Evans demonstrates that the criterion of pair wise convergence for all members of a panel is equivalent to the condition that the difference between the individual series, y_{it}, and the mean value for the series across all members of the panel at each point in time ($\bar{y}_t = \frac{1}{N}\sum_{i=1}^{N} y_{it}$) is stationary. Thus, the condition states that all members of the panel converge pair-wise if ($y_{it} - \bar{y}_t$) is stationary for each member i ($i = 1,2,..., N$) of the panel. Consequently, the null hypothesis of convergence can be interpreted as the unit root null in the panel unit root test (Evans, 1998).

As has been discussed above, in this context whether or not convergence is occurring can be evaluated by asking whether or not the autoregressive parameter β_i is zero for the panel data regression given by:

$$\Delta(y_{it} - \bar{y}_t) = \mu_i + \beta_i (y_{i,t-1} - \bar{y}_{t-1}) + \sum_{k=1}^{K_i} \phi_{i,k} \Delta(y_{i,t-k} - \bar{y}_{t-k}) + \varepsilon_{it} \qquad (7.8)$$

Where:

i (i = 1, 2, 3,….,N) is the number of observations (number of provinces in our sample); *t (t = 1, 2, 3,….,T)* is the number of periods (years) or sample size under the study as defined in (7.1); μ_i is fixed effects representing the individual province's average sample difference from the group mean ($y_{it} - \bar{y}_t$); *K (k = 1, 2, 3,…, Ki)* is number of lags in our study, and ε_{it} is error terms of province i in time t.

According to Evans, this specification is an augmented Dickey-Fuller (ADF) regression applied to the panel of income differentials between the individual provinces and the mean income value of the provinces as a group. In this case, the μ_i fixed effects represent the individual province's average sample difference from the group mean ($y_{it} - \bar{y}_t$), which is permitted to vary by province. The autoregressive parameter for the income differentials, β_i, becomes the key coefficient for determining the presence or absence of convergence, the lagged difference terms are intended to capture higher order serial correlation in the time series process for income differentials and the number of lags, *Ki*, are chosen in a manner to ensure that the remaining error terms ε_{it} are serial uncorrelated. Under this specification, the null hypothesis of the panel unit root is stated as

$$H_0 : \beta_i = 0 \text{ for all } i \quad (7.9)$$

against the alternative hypothesis

$$H_1 : \beta_i < 0 \text{ for some } i \quad (7.10)$$

In case the null hypothesis is rejected, it means that at least some subsets of the members of the panel are converging toward one another. In contrast, if the null hypothesis is not able to be rejected, it means that no subset of the members of the panel is converging toward one another.

To test this hypothesis, we employ the panel unit root tests used by Im, Pesaran, and Shin (2003). The tests used by Im, Pesaran, and Shin (IPS tests) (2003) allow the value for the autoregressive coefficient, β_i, under the alternative hypothesis to vary across provinces. On the contrary, the panel unit root tests by Levin and Lin (2002) employed in the first generation of panel unit root convergence tests such as Evans (1998) and Evans and Karras (1996) requires the autoregressive coefficient to be homogeneous under the alternative hypothesis, so that $\beta_i = \beta < 0$. Thus, more recent tests provide us with the additional flexibility of allowing the convergence dynamics to differ across provinces under the alternative hypothesis. Practically, the IPS tests are constructed on the basis of averaging the unit root tests for the individuals to produce a group mean test statistic. In one such test, Im, Pesaran, and Shin recommend constructing a *t*-bar statistic, which is based on averaging the individual augmented Dickey-Fuller unit root t-tests.

In our context, the procedure of the test is as follows. To compute the statistics, we first need to estimate the augmented Dickey-Fuller regression presented in equation (7.8) individually for each of the *i* (*i = 1, 2,.....,61)* provinces of the panel and then construct the 61 corresponding ADF *t*-statistics*,* t_i. These individual statistics are averaged to obtain the *t*-bar statistic ($\bar{t} = \frac{1}{N}\sum_{i=1}^{N} t_i$). Finally, because the distribution of the individual ADF *t*-statistics are not centered around zero under the unit root null hypothesis, it becomes necessary to adjust for this feature to ensure that the distribution of the t-bar statistic does not diverge under the null hypothesis as the number of individual members of the panel, *N* gets large. Luckily, under the null hypothesis, the mean of the individual t_i is a known constant as the sample size T get large, as the standard deviation of the individual t_i. Consequently, t-bar statistic is adjusted by subtracting off the mean and dividing by the standard deviation, so that the statistic becomes

$$\bar{z} = \sqrt{N}\left(\bar{t} - u\right)/ s \quad (7.11)$$

where *u* is the mean of the individual ADF *t*-statistic distribution, and s is the standard deviation of the individual ADF *t*-statistic distribution. Provided that the individual statistics are independent, then as Im, Pesaran and Shin demonstrate, this statistic will be distributed as standard normal under the null hypothesis. Thus, it will diverge to negative infinity under the alternative hypothesis so that large negative values can be taken to reject the null hypothesis.

7.3.3. The role of space and spatial autocorrelation

In case the Moran's I test and LM test developed in section 7.3.2 above fail to reject the spatial dependence among provinces, we go on to explore spatial effects or the role of space in the patterns of economic growth of the Vietnam's economy. The exploration of the spatial interaction is done based on the appropriate specifications whether it is spatial lag or spatial error depending on the specification tests based on the coefficient of the Moran's I statistics, and both the Lagrange Multipliers and Robust Lagrange Multipliers.

As our data is collected locally, so it is associated with spatial units. Thus, the data used for analysis is the spatial data. When a value observed in one location depends on the values observed at neighboring locations, there is the presence of a spatial dependence (Anselin, 1998, 2001). Spatial data may show spatial dependence in the variables and error terms. To consider this, in this section we introduce the spatial effect namely spatial autocorrelation. The spatial autocorrelation refers to the fact that one observation associated with a location depends on other observations in other locations. The analysis of the spatial autocorrelation is aimed at detecting spatial dependence among observations in the analysis of the *β-convergence.*

According to Anselin and Bera (1998, 241), the definition of spatial autocorrelation is as follows: "Spatial autocorrelation can be loosely defined as the coincidence of value similarity with locational similarity". In other words, the observation of a random variable in a given localization is partly determined by the observation of this variable in neighboring localizations (the neighborhood being defined by the spatial weight matrix). The absence of spatial autocorrelation is defined by random spatial distribution of the variable of interest.

The presence of spatial autocorrelation provides further information with respect to traditional statistics like mean or standard errors, because it gives an idea about the geographical distribution of the values of the variable studied. Furthermore, modeling spatial autocorrelation allows a researcher to take the existence, influence and size of geographic spillover effects into consideration. Before discussing about the spatial econometric models for spatial interaction, it is important to specify the modeling of interactions among regions. As we use N observations, it is necessary to impose a structure to spatial interactions, which is given by spatial weight matrix, W. The objective of structuring this matrix is to consider a neighborhood relationship for each region. There are some approaches used to construct weight matrix such as transport time between regions, distance between regions or binary concept. In this study we use a binary weight matrix which is defined as 1 if two regions share a border and 0 otherwise (Le Sage, 2001). We therefore have constructed a binary weight matrix where the connection relation regressor in the form of a spatially lagged dependent variable or in the error structure. The former is referred to as a spatial lag model (SAR[60]) and is appropriate when the focus of interest is the assessment of the existence and strength of spatial interaction. Spatial dependence in the regression disturbance term, or a spatial error model (SEM[61]) is referred to as nuisance dependence. This is appropriate when the concerns is with correcting for the potentially biasing influence of the spatial autocorrelation, due to the use of spatial data (irrespective of whether the model of interest is spatial or not). These two types of spatial dependent models sometimes jointly are present in the same regression. More precisely, spatial lag is present when the dependent variable in a place is affected by independent variables in that place and adjacent place, while spatial error occurs when the error terms across different units are correlated with each others.

[60] SAR stands for Spatial Autoregressive model.
[61] SEM denotes Spatial Error model.

The model from which spatial models can be derived is the classical linear regression model of the form:

$$y = \beta X + \varepsilon \qquad (7.12)$$

where: y is the $(N,1)$ vector of observation of the dependent variable, X is the (N, K) matrix of observations of the K explanatory variables, β is the $(K,1)$ vector of the unknown parameters to be estimated, and ε is the $(N,1)$ vector of the error terms, ε is $iid\ (0,\sigma^2)$.

The SAR model consists in inserting the endogenous spatial lag term, Wy in the set of explanatory variables. The spatial lag term Wy is correlated with the disturbances even when these disturbances are i.i.d. Formally, the linear model in (7.12) is transformed to be SAR as follows:

$$y = \beta X + \varphi Wy + \varepsilon \qquad (7.13)$$

where: Wy is the endogenous spatial lag term, and φ is the spatial autoregressive parameter expressing the interaction intensity between observations of the dependent variables. When spatial weight matrix W is standardized, $(Wy)_i$ is a weighted average of observations of regions in the neighborhood of region i.

The introduction of the endogenous spatial lag in the regression allows the assessment of the spatial dependence degree when the effect of other variables is controlled for. In order to see the two effects induced by this specification, it is worthy to write (7.13) in its reduced form:

$$y = (1-\varphi W)^{-1}\beta X + (1-\varphi W)^{-1}\varepsilon \qquad (7.14)$$

From the reduced form in (7.14), it is to note that the matrix $(1\text{-}\varphi W)^{-1}$ must be non-singular[62]. From the equation (7.14) we notice that the spatially lagged variable Wy is correlated with the error term ε. Thus, the OLS estimators would be biased and inconsistent. The simultaneity embedded in the Wy term must be explicitly accounted for in a maximum likelihood estimation framework (see Ord, 1975)[63].

The reduced form in equation (7.14) can be written in the inverse spatial transformation under the geometric expansion form as follows:

$$y = (1+\varphi W+\varphi^2W^2+\varphi^3W^3+.....)\beta X + ((1+\varphi W+\varphi^2W^2+\varphi^3W^3+.....)\varepsilon \quad (7.15)$$

The first term in the right hand side of the equation (7.15) presents the spatial multiplier effect, which means that the expected value of y in region i does not only depend on the value of explanatory variables in this region, but also on the value of those variables in all regions under consideration. From (7.15) we observe that this multiplier effect is decreasing over the distance. The second effect is the spatial diffusion effect, which is described by the second term in the right hand side of (7.15). This effect indicates that a random shock hitting a given region will gradually affect all regions under consideration. This effect is also decreasing over the distance.

Another way to deal with spatial dependence effects is to introduce an exogenous spatial lag variable (Rey and Montouri, 1999). In this model, spatial dependence in the

[62] This condition is satisfied if absolute value of $\varphi \neq 0$ and when $1/\varphi$ is not an eigenvalue of W.
[63] For more details, please refer to Anselin (1988), Anselin and Bera (1998) and Anselin (2001).

process of regional convergence can be incorporated by a spatially lagged explanatory variable *Wx*:

$$y = \beta X + \eta Wx + \varepsilon \quad (7.16)$$

For this specification, because the spatially lagged explanatory variable is exogenous, estimation of the spatial cross-regressive model can be based on the OLS method.

In addition to spatial lag, spatial autocorrelation may be present under another form namely spatially autocorrelated errors (SEM). This type of autocorrelation is preferred when autocorrelation is viewed more as a nuisance than a substantial parameter (see Florax and Nijkamp, 2004). Formally, the SEM model is as follows:

$$\begin{aligned} y &= \beta X + \varepsilon \\ \varepsilon &= \varphi W\varepsilon + u \end{aligned} \quad (7.17)$$

where φ is the spatial autoregressive coefficient and u is the vector of errors with usual characteristics . The error u is independent and identically distributed with mean zero and variance σ^2. The parameter φ reflects the intensity of the interdependence between residuals. According to Anselin (2001) due to non-spherical errors, OLS estimators are inefficient and the maximum likelihood method or the generalized method of moments is preferred.

Substituting the second equation into the first one in the (7.17) we get the reduced form of the SEM model:

$$y = \beta X + (1 - \varphi W)^{-1} u \quad (7.18)$$

Similar to the SAR model discussed above, the matrix $(1\text{-}\varphi W)^{-1}$ must be non-singular. From the equation (7.18) we observe that SEM model only includes a spatial diffusion effect as the term βX is not pre-multiplied by the inverse spatial transformation.

The process of the estimation is as follows. First we will estimate β by means of the OLS regression of both the unconditional and conditional models presented in equations (7.1) and (7.2) respectively. Based on the results, the speed of convergence and *half-life* presented respectively in equation (7.3) and (7.4) will be computed. Based on the diagnostics for spatial effects, spatial dependence models will be implemented using the maximum likelihood method.

7.4. Results

This section reports the empirical results and tests for convergence that developed theoretically and empirically, and then explores the role of space in the patterns of economic growth based on the models given in section 7.3.

- ***Absolute β-convergence***:

For the absolute *β-convergence*, we will apply the tools of spatial econometrics developed by (among others) Cliff and Ord (1973), Anselin (1988) and Anselin and Florax (1995) to test for unconditional *β-convergence* (absolute *β-convergence)* expressed in equation (7.1). The exercise bears on the provinces of Vietnam for the period 1990-2006. The reason that leads us to work on unconditional *β-convergence* is the lack of information on potential control variables presented in equation (7.2) for the case of conditional *β-convergence*.

First, we estimate the model (7.1) for analysis of the unconditional β-convergence using OLS. Taking variables of interest, the model to be estimated becomes:

$$y_T = \alpha + \beta y_{1990} + \varepsilon$$

with y_T is the vector of the average annual rate of per capita GDP growth in logarithms for each province between 1990 and 2006, T=16, y_{1990} is the vector containing the observations of the per capita GDP in logarithms in 1990, α and β are unknown parameters to be estimated and ε is the vector term with usual properties. The estimation results and the Moran's statistic (equation 7.5b) as well as LM tests (equation 7.6 and 7.7) are summarized in Table 7.2 below.

Table 7.2. OLS results of the absolute β-convergence and diagnostic tests for spatial autocorrelation

Dependent variable: y_T		Tests for spatial dependence	
R^2 adjusted	0.5074	*Moran's I (error)*	0.382 (0.702)
AIC	-380.882	*LM_{ERR}*	0.487 (0.485)
s^2	0.2611	*LM_{LAG}*	0.000 (0.987)
α	-2.1956* (0.000)	*RLM_{ERR}*	0.492 (0.483)
β	1.7965* (0.000)	*RLM_{LAG}*	0.005 (0.942)

Note: * indicates high significance level; numbers in parentheses are the p-values; Moran's *I* is the Moran test for spatial autocorrelation. LM_{EER} and LM_{LAG} are the Lagrange multiplier statistics which tests for the presence of spatial autocorrelation in the errors and a endogenous spatial lag respectively; RLM_{EER} and RLM_{LAG} are their robust counterparts; and AIC stands Akaike information criterion.

The results from the estimation of model (7.1) presented in Table 7.2 show that the estimated *β* parameter is positive (1.7965) and highly significant with high overall fit (adjusted R^2 is above 0.50 and high absolute AIC), which means that an absolute convergence process is not present in this studied sample. We therefore will not compute the convergence speed[64] and the half-life[65]. However, the rejection of beta-convergence hypothesis in Table 7.2 can contribute to the exploratory analysis presented in Figure 7.2 and 7.3 above which imply that there is a trend of divergence in regional income in Vietnam over times. In fact, income differentials between provinces based on Figure 7.2

[64] The convergence speed, ζ, is computed using equation (7.3)
[65] Half-life τ is calculated from the expression in equation (7.4)

and 7.3 can be stable in some regions or that divergence may occur in other regions. However, these figures are sketched based on a log scale of income. But if we use the absolute value of income, the pattern may be different (we don't make this comparison in this thesis).

Given the economic intuition that provinces are correlated by factors like international trade or knowledge spillovers, we have checked the presence of spatial autocorrelation. The most commonly applied statistic to test the presence of spatial dependence is the Moran's I. If significant, the sample is not randomly distributed and spatial autocorrelation is present among observations. However, this test does not specify the nature of dependence. In the case where it is significant, we first have to try to include more exogenous variables. If it is not possible because of a lack of these exogenous variables in this study, we must perform specification tests which will help us to determine the nature of spatial dependence.

The right-hand side of Table 7.2 reports the results of a number of diagnostics for the spatial dependence. Three different tests for spatial dependence are included, a Moran's I test, two Lagrange Multiplier Tests (*LM – error and lag*) and two Robust Lagrange Multiplier Tests (*RLM – error and lag*). According to a study by Anselin and Rey (1991), the first test is very powerful against both forms of spatial dependence: spatial lag and spatial error autocorrelation. Unfortunately, it does not allow for the separation between these two forms of misspecification. In contrast, the robust tests have displayed good power against a specific alternative in an extensive set of Monte Carlo experiments (Anselin *et al*., 1996). In our sample, all tests include Moran's I and two types of Lagrange Multiplier tests are allowed us to reject any presence of spatial dependence because all the tests are not significant with very high p-values. We therefore rule out of any spatial effects among observation in our sample.

- ***ADF (IPS) test for unit root***

The test above further confirms our suggestion given in the section on the informal data analysis of the income convergence which indicates that there was no sight of the per capita income convergence among provinces in Vietnam as all the tests provided in Table 7.2 reject the null hypothesis of the per capita GDP convergence in our sample. To further consolidate this conclusion, we next employ Evans' approach (developed in the latter part of section 7.3.1 above) to further testing for the per capita income convergence among regions in Vietnam using the augmented Dickey-Fuller test. To do so, the IPS t-bar test is applied for the entire sample size, from 1990 to 2006. Table 7.3 reports the results of the test. In this exercise, the lag truncation for the individual ADF unit root regressions were allowed to vary by individual province for both individual tests as well as the panel based test. In each case, the length of the lag was selected by a standard data dependent step down procedure, which is typically implemented for the ADF unit root test in time series regressions. The step down procedure involves starting with a sufficiently large number of lags and then sequentially eliminating the highest order lags one at a time until one of the tests is significant. In our study, we allowed this procedure to choose a different lag truncation for each province. For the arbitrary initial starting value, we rounded off to the nearest integer of the one fifth of the sample length. It means, in our sample size with T=17, that we started with a highest lag of three and then allowed the automated data dependent procedure to choose an actual number of fitted lags, which then varied between zero and three. Because the test with different lags was not significant, we report only the tables with results for the case with maximum truncation of three.

Consider first the results of the tests for individual province reported in Table 7.3. The first column to the right of the province name gives the value for the individual ADF t-statistic for the particular province computed using equation (7.11). The next column reports the associated marginal significant level known as 'p-value' for the given ADF t-statistic. Based on the p-values, we see that only three provinces (as Ha Giang, Cao Bang and Bac Kan) are able to reject the unit root null at 5% significant level or better, the remaining provinces are not able (even not being close to rejection) to reject the null of unit root as the statistics provided are far from rejecting the null hypothesis. This points to likelihood that on balance as discussed earlier, the majority of the provinces are not converging to one another since we cannot reject the panel unit root null for the differences ($y_{it} - \bar{y}_t$). On the other hand, the fact that a small subset of provinces provides rejections leads us to consider the possibility that there may be small subgroups that may be present with income convergence. In addition, even the combined evidence does not reject the null hypothesis of unit root as *t-bar* statistic for the panel test reflects this as ADF test is not significant. The last row of Table 7.3 gives a panel IPS t-bar statistics of –0.548 with p-value of 0.981. Thus, again we state, based on both the absolute *β-convergence* Moran' I test and the unit root test for co-integrated, that there is no sign of global convergence in per capita GDP in Vietnam.

Consequently, we consider the possibility that the apparent absence of convergence nationally can be attributed to the fact that at least some subsets of provinces are converging to separate regional clubs or that there is a presence of local convergence or any convergence club in the country. We are aware that the results for the full sample of provinces already indicate that the global convergence possibility is unlikely, since we cannot reject the likelihood that there is no sizeable subset of provinces which converge pair-wise within the sample. Given the presence of sampling variation and the fact that we cannot say a priori what constitutes a sufficiently sizeable subset on the basis of the full sample results, it is worth further investigating the convergence properties of candidate subsets of provinces. For example, we may propose that differences in geography or differences in preferential open-door policies at the provincial level may generate convergence club among the provinces of Vietnam. This may lead to a distinction as high income group versus low income group or coastal versus interior geographic regions, and that there may be present of the local convergence among these indentified groupings.

Table 7.3. Individual provincial ADF tests for unit roots

Province	ADF	p-value	Province	ADF	p-value
Hanoi	-1.056	0.936	Da Nang	-1.335	0.879
Hai Phong	-0.460	0.985	Quang Nam	-0.859	0.960
Vinh Phuc	-0.728	0.971	Quang Ngai	-1.530	0.819
Ha Tay	-0.674	0.975	Binh Dinh	-1.196	0.911
Bac Ninh	-0.989	0.946	Phu Yen	-1.258	0.898
Hai Duong	-1.445	0.847	Khanh Hoa	-1.653	0.771
Hung Yen	-2.138	0.525	Kon Tum	-2.384	0.388
Ha Nam	-1.191	0.912	Gia Lai	-1.533	0.817
Nam Dinh	-0.201	0.991	Dak Lak	-0.548	0.984
Thai Binh	-1.598	0.793	Lam Dong	-1.517	0.823
Ninh Binh	-0.454	0.985	Ho Chi Minh City	-2.398	0.381
Ha Giang	-8.244	0.000	Ninh Thuan	-0.797	0.966
Cao Bang	-3.487	0.041	Binh Phuoc	-0.409	0.987
Lao Cai	-0.694	0.974	Tay Ninh	-1.496	0.831
Bac Kan	-3.911	0.017	Binh Duong	-2.289	0.440
Lang Son	-1.769	0.719	Dong Nai	-2.282	0.441
Tuyen Quang	-1.075	0.933	Binh Thuan	-0.172	0.992
Yen Bai	-0.990	0.945	Ba Ria-Vung Tau	-1.080	0.932
Thai Nguyen	-1.126	0.925	Long An	-0.824	0.964
Phu Tho	-1.288	0.891	Dong Thap	-1.565	0.806
Bac Giang	-0.948	0.951	An Giang	-2.063	0.567
Quang Ninh	-1.242	0.901	Tien Giang	-1.690	0.755
Lai Chau	-2.044	0.577	Vinh Long	-0.871	0.959
Son La	-1.567	0.805	Ben Tre	-2.477	0.339
Hoa Binh	0.785	1.000	Kien Giang	-1.876	0.667
Thanh Hoa	-0.510	0.983	Can Tho	-1.424	0.854
Nghe An	-1.061	0.935	Tra Vinh	-2.175	0.504
Ha Tinh	-1.888	0.661	Soc Trang	-2.504	0.326
Quang Binh	-1.193	0.912	Bac Lieu	-1.809	0.701
Quang Tri	-0.436	0.986	Ca Mau	-1.933	0.637
Thua Thien-Hue	-1.287	0.891	**ADF**		**p-value**
Panel IPS-t bar			*-0.548*		*0.981*

We now examine various regional sub-groupings for the possibility of geographically-based convergence clubs. We design the high income group versus the Central Coastal group, the North interior group and the South interior group. The last two groups are present for the interior groupings. As described earlier in section 7.2, our sample of high income subgroup consists of six provinces having highest per capita GDP level. The Central Coastal subgroup consists of six coastal provinces excluding the relatively high income provinces such as Da Nang, Khanh Hoa and Phu Yen, etc. The North interior subgroups consists of five low income provinces in the Northern mountainous region, and Southern interior subgroup consists of seven low income provinces in the Southeastern region and the Mekong River Delta exclude coastal and high income provinces in these regions. Table 7.4 reports group results of ADF test for these groupings for both each province individually and for the subgroup as a whole (panel test) in the last row of Table 7.4. Notice that, statistically, the individual test values for the same province tend to differ depending on the grouping in which it is included (in Table 7.4) compared to the whole country as a sole group (in Table 7.3). This difference is because when we test ($y_{it} - \bar{y}_t$), the values for $\bar{y}_t$ differ between groups compared to the whole country. Looking at the results of the ADF test for each province individually, in all groups we were able to reject the null hypothesis that all of the ($y_{it} - \bar{y}_t$) are non-stationary for one province in each group. In the group of high income provinces, the ADF test for unit root was able to reject the null hypothesis for the case of Da Nang at 0.1 significant level with p-value of 0.006. In the Central Coastal group, IPS test was able to reject the null of unit root for the case of Quang Nam province at almost all significant level (p-value of 0.000). The IPS test was also able to reject the null hypothesis of unit root for the case of Son La province in Northern interior group and Dong Thap province in the Southern interior group (respective p-value of 0.004 and 0.000). We therefore conclude that there is no presence of any convergence club in the country. This finding is not in line with our earlier presumed presence of a separate high income convergence club or central coastal convergence club or northern interior convergence club as discussed above in the informal data analysis of the possible income convergence or divergence process in Vietnam. Indeed, there may be a possible another pattern that divergence may be taken place given the size of beta and the high statistical significance. The last row of Table 7.4 provides panel IPS t-bar for the ADF test for convergence clubs (convergence subgroups). The panel IPS test has failed to reject the null hypothesis of unit root for any groupings. The panel IPS t-bar tests for the group of high income provinces has a high p-value of 0.258, for the group of central coastal provinces has a high p-value of 1.000, for the group of the Northern interior provinces has a high p-value of 0.808, and for the Southern interior provinces has a p-value of 0.189. Thus, there is no sign of the per capita GDP converging in any groupings of provinces of Vietnam.

Based on the results of the absolute *β-convergence* tests for convergence in per capita income and ADF tests (IPS tests) for unit root (evidence for the convergence in per capita income), the results illustrate an interesting picture for the growth patterns among provinces of Vietnam during the *doi moi* period. As a general phenomenon commonly happens in the developing economies, per capita GDP among the provinces do not appear to be converging. When you have tested that there is no convergence, this does not mean that there is growing disparities in GDP per capita over the provinces of Vietnam. As examples of grouping convergence investigation above, it does not appear to be the case that this can be explained by the presence of a simple, dual-convergence club that distinguishes between high income provinces and low income provinces or between central coastal provinces and interior provinces. In all cases as the above tests indicated per capita incomes do not appear to be converging toward one another regardless of which groupings we consider.

Table 7.4. The ADF Test for local convergence among geographic sub-groupings

Province	High income subgroup		Central coast group minus high income		Northern interior		Southern interior	
	ADF	p-value	ADF	p-value	ADF	p-value	ADF	p-value
Hanoi	-2.909	0.159	-	-	-	-	-	-
Hai Phong	-1.040	0.938	-	-	-	-	-	-
Da Nang	-4.102	0.006	-	-	-	-	-	-
Ho Chi Minh	-1.695	0.753	-	-	-	-	-	-
Vung Tau	-1.209	0.908	-	-	-	-	-	-
Binh Duong	-1.780	0.714	-	-	-	-	-	-
Thanh Hoa	-	-	-1.583	0.798	-	-	-	-
Ha Tinh	-	-	-2.932	0.152	-	-	-	-
Quang Binh	-	-	-1.956	0.625	-	-	-	-
Quang Tri	-	-	-1.879	0.665	-	-	-	-
Quang Nam	-	-	-5.012	0.000	-	-	-	-
Quang Ngai	-	-	-1.761	0.723	-	-	-	-
Ha Giang	-	-	-	-	-2.320	0.423	-	-
Cao Bang	-	-	-	-	-0.267	0.990	-	-
Bac Kan	-	-	-	-	0.335	0.988	-	-
Lai Chau	-	-	-	-	-0.134	0.992	-	-
Son La	-	-	-	-	-4.258	0.004	-	-
Ninh Thuan	-	-	-	-	-	-	-2.185	0.498
Binh Phuoc	-	-	-	-	-	-	2.127	1.000
Binh Thuan	-	-	-	-	-	-	0.087	0.995
Dong Thap	-	-	-	-	-	-	-9.886	0.000
An Giang	-	-	-	-	-	-	0.253	0.996
Tien Giang	-	-	-	-	-	-	0.693	0.997
Vinh Long	-	-	-	-	-	-	-0.910	0.955
Panel IPS-tbar	*-2.649*	*0.258*	*1.234*	*1.000*	*-1.558*	*0.808*	*-2.822*	*0.189*

7.5. Conclusion and discussion

In chapter seven we have developed the theoretical framework of and carried out the two different kinds of empirical tests (absolute *β-convergence* test and ADF panel IPS t-bar test) for the convergence in per capita GDP, and Moran's I and the both the LM tests (Lagrange Multiplier test and Robust Lagrange Multiplier test) for spatial dependence among provinces of Vietnam during the *doi moi* period using data provided by the General Statistical Office of Vietnam for a period with a relatively high rate of economic growth stretching from 1990 to 2006. This exercise presents an important case study for examining regional income differences that accompany rapid economic growth. By using recent econometric developments which account for the time series properties of the data for the IPS t-bar test, our primary results confirm that there is no sign of a converging pattern in per capita GDP among provinces of Vietnam both nationally and locally (groupings or clubbing). Since 1990, the per capita income in the majority of provinces appears to be diverging away from each other. Even though, however, the tests for convergence are all rejected, we are not in a firm stand to conclude that the regional income level in the country is really diverging. Instead, based on the Figure 7.2 and 7.3, informal exploratory analysis and concrete elaboration of the relevant literature, we may say that the divergence may be taken place given the size of beta and the high statistical significance.

In parallel with the test for the per capita GDP convergence, the chapter did also attempt to test for the spatial dependence among provinces in order to explain the growth pattern of one province in relation with that of others. However, the Moran's I and both the LM and Robust LM statistics given in Table 7.2 did not support our prior assumption of the spatial dependence among provinces in the growth progress.

These finding results point to the idea that these features of Vietnam's experiences may be important in explaining this systematic and persistent regional income divergence in other developing economies in their initial stage of development and growth. In our tests for the income convergence among regions in Vietnam, we found no glimpse of convergence process nationally (in which Vietnam is considered as a whole), but there are very little evidences of club convergence occurred among smaller groupings. These findings match with both the finding by Kuznets (1955) which says that income is diverging in the initial period of development, and the finding by Krugman, in which he indicates the importance of the location of economic activities or geographical effects (Krugman, 1991). The study also finds support for the findings about the pattern of income convergence by studies for developing and transitional economies like Vietnam, such as China and Russia by, among others, Pedroni and Yao (2005), Zhang *et al.* (2001), Bradshaw and Vartapetov (2003), Chen and Fleisher (1996), Demurger *et al.* (2002), and Weeks and Yao (2003), which found evidence for both the convergence and divergence in regional income in these countries. However, these studies for the cases of China and Russia did not find the evidence for global (or national convergence). Instead, they found that some club convergences existed in these countries, but generally, there was no evidence of global (national) convergence found for the regional growth pattern of income levels these countries.

Historically, experience elsewhere in the world suggests that few countries have succeeded in maintaining political stability under conditions of severe disparity. Despite this fact, Vietnam has never gone through such a circumstance. However, the existence of regional income disparities is of considerable interest as it bears directly on the sustainability of economic reform (*doi moi*) and open-door policy. The findings from this analysis add to the growing body of literature that further confirms a serious potential risk if this disparity is not addressed that growing regional income inequality might harm the stability of Vietnam economic reform process and slow down the future growth.

The findings also have important implications for the countries in the Southeast Asian region and in other regions of the world as well as transitional economies that show a similarity in economic growth path with Vietnam's. It is also a good implication for other countries in other regions that are contemplating prospects for greater economic integration and cooperation through the channel of regionalization and globalization. These findings also help other regional economic groups in developing regions as in Africa, and Latin America to understand the fact that more rapid growth associated with greater openness to trade in goods and services may also be associated with larger regional income differences.

Chapter eight
Overview, conclusions, discussions and recommendations

8.1. Overview of the study

This study has examined the size of and changes in multidimensional regional inequality over the period from 1998 to 2004 for education and health, and changes in regional inequality in per capita GDP for over nearly two decades, starting from 1990 to 2006 in Vietnam after the country embarked its economic reform policy in the late 1980s. The study has also analyzed the relationship between regional inequality in per capita GDP and regional economic growth, patterns of growth and the space effects in regional income convergence, as well as contiguity effects, and the levels and trends in between- and within-region welfare disparities. This study particularly seeks to answer the following research questions:

- What is the size of regional multidimensional inequality in health and education?
- How has the regional multidimensional inequality in health and education evolved?
- What does the regional profile based on the rankings of composite index with respect to education and health look like?
- What is the role of contiguity effects in the health status of the people or does a spatial dependence among health status and health facilities in the country exist?
- What are the magnitudes and trends of regional inequality in per capita GDP?
- What is the relationship between regional inequality in per capita GDP and regional growth in Vietnam? What is the role of space in this relationship?
- What is the regional pattern of economic growth? Does the regional per capita GDP converge to or diverge from one another?

To resolve the above-mentioned research questions and to attain the research's objectives, several methodological approaches have been employed and discussed. A method for regional welfare inequality and the detailed procedure for analyzing regional inequality has been elaborated and developed. The approach to analyze the relationship between regional inequality in per capita GDP and regional growth has also been presented. Lastly, several tests for convergence as well as the role of space have been developed and carried out in conjunction with growth patterns over nearly two decades, spanning 1990 to 2006. This chapter is set to summarize the main findings of the present study, to discuss them with the relevant literature and to give recommendations.

8.2. Conclusions

The concept of welfare is somehow not fully and completely understood by most people in Vietnam. Before implementation of the economic reforms (*doi moi)* in 1986, the Vietnamese people were equally able to access almost all services provided by the government, irrespective of religion, ethnicity or social status. Like other countries in the communist system, distribution of welfare in Vietnam was relatively equal in the past, although the quality of the system with respect to services offered was still questionable in terms of catching up to the international standard. Before embarkment of the *doi moi* in 1986, most Vietnamese people enjoyed a free healthcare system and pupils and students had free access to education at all levels. These free systems culminated in good health records attained by the country, high enrolment rates and a high rate of literacy among the Vietnamese people, compared with developing countries with the same or slightly higher level of development. The high economic growth rate at on average 6-9% annually during the *doi moi* has affected macro-economic performance in the country, and as a result, has affected the welfare system of the country, as the explanatory analysis using the Gini coefficient, proved that regional inequality in the country has risen ever since.

The measurement and analysis of regional welfare inequality for health facilities and health status over the 1998-2004 period for three separate samples show that total regional inequality has different patterns for healthcare facilities and health status. Generally, regional inequality in both the healthcare facilities and health situation moderately decreased in Vietnam over the 1998 – 2004 period. However, there was a sign of increase during this stretch of time, as calculated Theil's second measure presented in chapter four show a slight increase in inequality in both components of health over the 1998-2002 period, and then it decreased in the next period. This moderate increase and then decrease in the inequality of health facilities and health status can be attributed to the effect of economic growth and semi-privatization of the health system in Vietnam. High economic growth has made almost all people better off in all regions of the country. However, a rather small portion of the population in the wealthier regions gains from the growth much more than part of the population, who live in poorer and more backward regions. As a result, health facilities and health status in backward and remote areas have been improved. In addition, *doi moi* has improved the quality of intra-region transportation that reduces time travel between regions and provinces. This improvement propels people to seek better health services in other regions rather than in their own. All these factors would bridge the gaps in regional differences in health facilities and health status.

Referring to the importance of spillover effects, as the healthcare services seem to be more equally distributed in the case of contiguity than in the case of non-contiguity, since Theil's second measure in the former is lower than that in the latter. This finding indicates that if we ignore the existence of spillover effects, the regional inequality will be exaggerated. Therefore, the findings suggest that it is important to take the contiguity effects into consideration.

These were interesting findings and contrast with our initial expectation that the *doi moi* in Vietnam may have deteriorated the situation of healthcare facilities and health status between regions in the country, because in the past, people were equally able to get free access to health services. However, the *doi moi*, has made both the country and people wealthier in terms of real income. Thus, with the improvement of both personal income and regional income, more investment in basic infrastructure, including the health system, should be made in order to improve the overall living standard of people. As a result of income improvement, people will also spend more on their healthcare in order to improve their health status. Thus, the situation of health facilities and health status during and post doi moi period are better than in the *pre-doi moi* period.

Based on the composite indices with respect to health facilities of 61 provinces, we find that the Red River Delta records the highest level of the composite index in all three years, which implies that the region is most-favored with respect to healthcare facilities. The Central Highland was registered with the lowest composite index which means that the region is least-favored in term of healthcare facilities despite the fact that this region has received the highest share of budget recently. This can be explained as that, in the past the health facilities in the region were very poor and backward and in severe shortage. Recently high investments in the health sector accrued to this region just partly to compensate for such a shortage and to upgrade currently available health facilities of the region. Such investments could not make the health facilities of the region better than the ones of other regions with good health facilities such as the Red River Delta. If we look at the ranking order by regions, there is evidence showing the changing patterns in health indicators between the surveys. We observe that the ranking order of most regions do not change in the 1998-2006 period except that there are swaps between the North East and the North West, and between the South East and the South Central Coast.

The tests for spatial dependence among health components have been conducted for several specifications of the relationship between the composite index with respect to health status and the composite index with respect to health facilities, and between the composite index with respect to health status and the indicator of health facilities. The results of the test confirm the existence of the spatial effects between the two dimensions of health. The test also proves that the spatial error model is an appropriate specification for analysis of the relationship between the composite index with respect to health status and the composite index with respect to health facilities, while the spatial lagged model is a more appropriate specification to analyze the relationship between composite index with respect to health status and indicators of health facilities.

Generally, these findings are in conformity with the fact that there are spillover effects in health facilities that could impact on health status among regions in Vietnam. In addition, the research draws a conclusion as to the relation between healthcare facilities and health status, as the findings reveal that the regions with a high composite index with respect to health status do not completely coincide with those having a high composite index with respect to healthcare facilities. This means that regions which are better equipped with health facilities - in other words, the regions get more investment in the healthcare system or have a higher share of the budget spent on the healthcare system - may not yield a better health status for people living in those regions. The main reason for this result is that, as the statistics and findings reveal, the regions with more investment or a higher share of budget spent on the healthcare system are the ones with a backward infrastructure and a poor socio-economic level of development. These regions are mostly the mountainous, remote and backward regions. Furthermore, with the short time span of six years, the budgets spent on health may not have enough time to reap the positive change in people's health status (lag effect). In addition, in the mountainous and backward regions, bad living habits of the people, such as tobacco and narcotic smoking, consuming large volumes of alcohol, insufficient hours of sleep, overworking, less-nutritional meals and mental stress by poor and hard life, are considered unhealthy lifestyle that will negatively affect the health status of the people in the regions (GSO, various years), while most lowland regions with a lesser share of the budget spent on health already have good records as to health status.

The analysis based on the values of the indicators' coefficients of the education facilities reveals that there is a changing tendency in the influence of each indicator on the regional distribution of education facilities. The findings reveal that the variable number of primary education teachers is the most dominant variable and the number of higher education teachers is the least dominant, while other variables change their influence from year to year. These findings are in conformity with the practice in Vietnam, where more attention has been paid to improving quantity and quality of education in recent years. This is fruitful as the country has successfully provided a rather equal spread of secondary teachers across all regions and provinces.

For regional inequality in education facilities, the research shows that total inequality, and between-region term and within-region components all go into opposite directions between the two periods. However, the change in the within-region inequality is in contrast with the changes in total regional inequality and the between-region component, as the total regional inequality and the between-region component in education facilities first decreased during the period 1998-2002 and then increased during the period 2002-2004, while the within-region inequality increased during the 1998-2002 period, then decreased during the 2002-2004 period. As the research findings reveal, the distribution of education facilities among the regions in Vietnam is a bit more unequal over time. Meanwhile, the change over time of the within-region component of inequality is very moderate. This implies that the opportunities to access education facilities for people within a group are almost unchanged over time.

For regional inequality in enrolments, our research findings show that the pattern of total regional inequality and its decomposed terms are the same, as all of them increased in the first period from 1998 to 2002, and declined in the next period of the study between 2002 and 2004. These findings imply that enrolments in the year 2004 were more equal. This means that the enrolment situation has improved over time. The structural shares of between-region and within-region inequality in the total inequality tell us that in all three years the between-region term makes up a major part in the total inequality. This implies that the differences in enrolments between regions are more significant to the total inequality than the differences in enrolments within regions.

Referring to the comparative composite index with respect to education facilities and the composite index with respect to enrolments, the research findings show that there is a sharp contrast between the two dimensions of education, as the most-favored regions with respect to enrolments are those in the lowland areas with least- and medium-favored education facilities. The least-favored regions with respect to the education enrolments are those in the upland areas, which are very high-favored with respect to education facilities in terms of the composite index. It has been explained earlier in chapter five that high drop-out rates are found in the areas with a high incidence of poverty. The findings coincide with the fact that the improvements in education facilities do not result in the improvement in education enrolments in many regions in Vietnam (MOET and Vietnam Development Report, 2005).

The research findings on the measurement and analysis of the regional income inequality in Vietnam have confirmed that, generally seen, Theil's second index of the regional income inequality in the country moderately increases over time. The results of decomposition of the Theil's second index of regional income inequality indicate that, over time with the process of economic development, the distribution of income within the region becomes more and more equal compared with the initial period, while the distribution of income between regions becomes more unequal. However, the size of the within-region inequality is different among the regions in the country, as the within-region 's inequality shows different trends, as is revealed in the first part of chapter six.

By using different estimation techniques to explore the relationship between regional inequality and regional growth, we have found that there are different patterns in the relationship between them. The estimates from the panel data regression for both the entire growth period and for different growth episodes show a positive relationship between the rate of economic growth and the initial level of per capita income in all the cases studied. However, the relationship is statistically significant for only the case without both control variables and dummies in the growth model for the entire growth period estimation, and for the case of the estimate for the five-year growth episode, starting from 2001 and for the ten-year growth episode, starting from 1996. The analysis based on different growth episodes reveals that the entire growth episode seems to produce a positive relationship between the rate of growth and initial level of per capita income significantly and more consistently. Meanwhile, using pooled OLS technique we have found that there are negative relationships between growth and inequality for all the cases studied: without both controls and regional dummies, with the controls and no regional dummies, and with both controls and regional dummies. However, the relationship has been found to be significant only for the case with both controls and regional dummies.

The test for the validity of the method (either the fixed effects or random effects is more appropriate and preferred) indicates that the fixed effects method is appropriate and preferred for the cases: without both the controls and dummies and with controls but no dummies in the growth model, while the random effects technique is more appropriate and more preferred for the case with both the controls, regional and time dummies. Interestingly, estimates by both fixed effects and random effects methods reveal that there is a positive and negative relationship between regional growth rate and initial level of per capita income. The positive and not statistically significant relationship has been found in all cases when the model is estimated by using fixed effects technique; and in the cases without both the controls and dummies and with controls but no dummies when the model is estimated using random effects method. Meanwhile, the negative and not statistically significant relationship between regional inequality and regional growth of per capita income has been found for the case that the model is augmented with both the controls and dummies using the random effects method. The findings of the mixed relationship between regional inequality and regional growth by different methods of estimation are also in line with the findings from the analysis of regional inequality in both developed and developing countries (Perotti, 1996; Li and Zou, 1998; Forbes, 2000; and Barro, 2000).

The empirical results of the relationship between regional inequality and regional growth show that the estimated relationship between inequality and growth changes when using different estimation techniques. However, the estimation using panel data regression, the fixed effects and most cases of random effects methods seem to support the theoretical argument and empirical findings that there is a positive correlation between growth and inequality for the developing economy like Vietnam's, while pooled OLS regression yields the results of a negative relationship between the two variables: inequality and growth. The same results were found in many empirical studies (Alesina and Rodrik (1994), Persson and Tabellini (1994), Perotti (1996), Easterly (2001), Forbes (2000) and Barro (2000). Moreover, based on a statistical point of view and several criteria such as R^2, AIC and BIC indices, we conclude that the method of Pooled OLS regression is most appropriate and more preferred.

Lastly, the findings from the test for convergence of the regional per capita GDP in Vietnam by applying recently developed econometric techniques, the research's primary results confirm that there is not any sign of converging patterns in per capita GDP among provinces of Vietnam, both globally and locally (or groupings). However, the exploratory analysis in chapter seven implies that since 1990, the per capita income in the majority of provinces has appeared to be diverging from each other. Even though the tests for convergence are all rejected, we cannot firmly conclude that the regional income level in the country is really diverging. Instead, based on informal exploratory data analysis and detailed elaboration of the relevant literature, we may say that the divergence of regional income may take place in Vietnam, given the size of beta and the high statistical significance.

Parallel with the test for the per capita GDP convergence, the test for the spatial dependence among provinces in order to explain the growth pattern of one province in relation with that of others has been also conducted. However, the Moran's I and both the LM statistics do not support our prior assumption of the spatial dependence among provinces in the growth progress.

The findings from the analysis of the regional per capita GDP convergence also help other regional economic groupings in developing regions in Asia, Africa and Latin-America to become aware of the fact that more rapid growth associated with greater openness to trade in goods and services may also be associated with larger regional income differences within a country. This can be explained as follows. Under the effects of globalization and/or regionalization, the developing and poor countries in Asia, Africa or Latin-America are able to export more of their own goods (and to import more), and they can be expected to be recipients of foreign direct and portfolio investments from capital-rich countries. However, foreign direct and portfolio investments will not accrue equally among regions within a country. They will concentrate more in more favored and advantageous regions (mostly urban or nearby areas) with a good infrastructure, abundance of labor and favorable socio-economic conditions. As a result, it may produce movement of work force from the low wage sector (agriculture), where wage differentiation is minimal, to medium skill sectors (in urban areas), where wage differences are larger. Then, even if the ratio between the top and bottom decreases (that is, the ratio between high skill and low skill wages becomes smaller), overall wage or income difference may increase, simply because of the greater wage differentiation in the middle. This is called the negative effect of regionalization and globalization on income or distribution of welfare, by which the movements of labor, goods and services and resources freely move between regions in a country.

8.3. Discussion

The results of regional inequality in several welfare components found in the case study of Vietnam partly contribute to and consolidate the firm foundation in the literature of development economics, which states that welfare inequality increases in the initial stage of development, as found by Kuznets (1955), and recently proved by many contemporary empirical researchers with their studies on developing and transitional economies, such as Sen (various years); Heltberg (2003); Liu (2001); Lugo (2004); and Justino *et al.* (2004).

The study also draws a conclusion that there is both a positive and negative relationship between regional inequality and regional growth during the period of high growth rates of Vietnam's economy brought about by the implementation of the *doi moi* policy. The pattern of the relationships depends on the methods of estimation used in the analysis, whether it is panel data or pooled regressions, fixed effects or random effects techniques. In a sense, the findings are in line with the findings that there is a positive relationship between inequality and growth, using panel data regression, fixed effects and random effects estimation in studies by, among others, Alesina and Rodrik (1994), Perssion and Tabellini (1994) and Parotti (1996), while the findings of a negative relationship between regional inequality and regional growth using the pooled estimation technique that is supported by Forbes (2000) and Li and Zou (1998). In another sense, however, there is still doubt about the relationship between regional inequality and growth in the case of empirical analysis for Vietnam. The findings on the relationship do not really fit well in the existing literature on growth and inequality in developing economies, because different methods of estimation yielded different outcomes. Moreover, I am not sure that the results of this chapter may be strong enough to justify that the study finds empirical support for some specific channel linking inequality to growth, because the different results are obtained by different techniques of estimation. In addition, the length of the growth period of 17 years is not that short, but not long enough either, as I observed from the literature that the duration of the growth period used for analysis usually stretches over 30 years. Thus, the empirical findings would be more impressive and more convincing, if we had a set of growth data that stretches a longer time duration, say, around three decades or more.

In chapter seven, the study's findings also indicate that since 1990, per capita incomes in the majority of provinces have appeared to diverge from one another, even though we cannot firmly conclude that, despite the fact that the tests for the regional convergence have all been rejected. These findings point to the idea that these features of Vietnam's experiences may be important in explaining the systematic and persistent regional income divergence in other developing economies in their initial stage of development and growth. It is in line with most studies in the literature of growth and development and growth patterns for both developed, and developing and transitional economies done by various researchers among others, Zhang *et al*. (2001); Pedroni and Yao, (2005); Solanko (2003); and Bradshaw and Vartapetov (2003).

However, several tests for the role of space in the growth process do not support our prior assumption of the spatial dependence among provinces in the growth progress. This finding is different from many studies on the convergence process and the role of space for the developing and transitional countries (Azzoni, 2000 and Krugman, 1991). This is the reason why we did not go on to examine the spatial effects or role of space in growth patterns as our initial aim has set for.

8.4. Limitations of the study

Despite the significant findings from our research, however, there are several limitations in the approaches used to study regional inequality in welfare and per capita GDP, and the methods resorted to explore the relationship between the regional inequality in GDP per capita and the regional growth, as well as the approach to investigating the growth patterns in Vietnam. The biggest limitation lies in the number of welfare indicators chosen for the analysis. PCPC or PCA can be applied to as many variables as possible, but in our study, due to unavailability of the data on other important indicators, many important aspects of health and education are missing. This may affect the empirical analyses of our study. Another limitation lies in the length of the period used for measurement and analysis of welfare inequality, except income. The duration that stretches over 6 years, from 1998 to 2004, is considered to be short, compared with the studies conducted in the literature on welfare inequality, in which the length of the period studied is usually more than a decade. This rather short period may affect the empirical findings of the research in a way that it will not provide a complete panorama of the evolution of a welfare's inequality that targets on several policies, because such a rather short period may not have enough impact for the intervention policies on the welfare studied.

For the chapters six and seven, where the models have been augmented with several controls, dummy as well as conditional variables, there is still a lack of many important variables, due to unavailability of statistical data on those variables. This can lead to biased and inconsistent findings, because economic growth and development are affected not only by the factors included in the models, but also by many factors that are still missing in the models we use, such as public, private and foreign investment, employment and fiscal policy, trade (import and export). All these variables on one way or another affect the economic growth of the country.

Besides, data used for the analysis in the present study are based on several archived sources, including from GSO and the World Bank, most of them being from GSO. Thus, the combination of the data that come from several sources may have limitations. These constraints should always be borne in mind, especially when making comparisons. In addition, there is a problem, due to the combination of data coming from several sources. This may somehow lead to inaccurate analysis and biased findings, because the data is gathered, collected and processed differently by each organization. Different techniques of collecting and processing the data make them rather incompatible and incomparable. However, although we have used several methods to pool and combine different data sources into a set for analysis, we do not really expect the final data set used for analysis will be as good as a set of data produced by one unique organization, either GSO or the World Bank.

8.5. Recommendations for policy

Based on the research findings, the following recommendations for policy can be made.

For the health system, the findings from the analysis of health inequality suggest that the government should spend more to improve the infrastructure of the health sector, in order to meet the demand of using health facilities by people, because the study reveals that people are more willing to spend on keeping a good health by increasing nutrition to realize healthy and better daily diets, having health checkups more regularly and spending more on medication in case of disease, as statistics from the household expenditure surveys showed that household expenditure on health increased by almost 30% during the 1998-2004 period. In addition, findings in chapter four reveal that the within-region component of the regional inequality is larger than the between-region share, which means that the health facilities within a region are not distributed equally. That is why it is important that the government should make sure that the budget is spent equally on the health system, more healthcare equipment should be being installed for the poor, backward and remote areas, in order to close the gaps in within-inequality in health facilities and health status.

For the education system in Vietnam, the findings from the research reveal that the pro-poor policy culminated in education does not seem to work well in this case, because the opportunities for access to education facilities for the poor are almost unchanged, though it does not lead to more inequality in education facilities. In addition, more investments made in education in the poor, backward and remote areas do not result in better enrolment rates in these regions, due to poor and lack of labour force in the households living in the regions. Thus, the government should consider improving the economic conditions in these regions, in order to generate more income for households in the areas that would help them afford for their children to go to school. This is a very important channel for a long-term strategy of poverty eradication, because education has been considered for a longtime a powerful tool for poverty eradication (World's Economic forum, Davos-2006).

For the relationship between inequality and growth as analyzed in chapter six, historically, experience elsewhere in the world suggests that few countries have succeeded in maintaining political stability under conditions of severe disparity. Despite this fact, Vietnam has never gone through such a period. However, the existence of regional income disparities is of considerable interest, as it bears directly on the sustainability of economic reform (*doi moi*) and open-door policy. The findings from this chapter add to the growing body of literature that further confirms a serious potential risk, of a growing regional income inequality that might harm the stability of Vietnam's economic reform process and slow down the future's economic growth, is not addressed. The findings also have important implications for the countries in the Southeast Asian region and in other regions of the world that show a similar growth pattern to Vietnam. These findings also help other regional economic groups in developing regions such as in Africa and Latin-America to be aware that more rapid growth associated with greater openness to trade in goods and services may also be associated with larger regional income differences.

8.6. Suggestions for further research

Besides the very useful and interesting findings, the present study is still incomplete in some sense. Thus, there are some suggestions for further research, in order to make the topic more complete and more useful, and more interesting for the context of Vietnam in particular, and developing and transitional countries in general. Firstly, for welfare inequality analysis, it is more complete and useful, if the data is collected for a longer duration, say 20 to 30 years. In addition, when we use an aggregate function of welfare indicators for composite index computation process, it is much better if we have data on useful variables that are still absent from the list of indicators of the welfare situations under study. If the length of period is longer, the research findings should be more robust and consistent, and if the model is augmented with more welfare variables, the findings will cover more aspects of the welfare situations studied, which would make the analysis more complete, more robust and more consistent.

Besides, we have measured and analyzed different welfare components separately for Vietnam, but there is still a lack of systematic studies in which all the welfare components and income are combined into a composite welfare, in order to measure and analyze total inequality of welfare and calculate a composite index for welfare and income as an aggregate composite index. If this kind of study can be carried out, its findings will be very interesting and useful, because once we calculate an aggregate composite index and inequality of the combined welfare for Vietnam, we will have a complete picture of the situation, magnitude and evolution of welfare inequality in Vietnam. We can also draw a map of favored regions with respect to aggregate welfare composite index, which will be very useful for the policymaker to target on the less-favored regions in terms of welfare systems and development of welfare.

For the study of the relationship between inequality and growth, the study will be more complete, if we go a step further to determine the determinants of growth and determinants of inequality. Thus further research on this topic for Vietnam should include more control variables that have both effects on growth and inequality, and take the determinants of growth and inequality in Vietnam's economy into account.

Lastly, the test for the convergence process carried out in the last chapter of the thesis has been rejected, which means that there is no way of convergence in per capita GDP in Vietnam. However, the test for the role of space in the growth process has also been rejected, which means that there is no evidence of the spatial dependence in the process of economic development of Vietnam, at least not in the short and medium run (less than 20 years). This has raised doubts about the conclusion of no spatial effects in the process of economic development. So I suggest further in-depth research on the topic for a longer length of growth episodes that will help us to detect the role of space in economic development. Furthermore, if we have a set of data on growth and relevant variables that covers both the pre-reform (before 1986) and post-reform period, the research will be more complete and robust, and it will help us to make a reliable and significant comparison of growth patterns between the two episodes of growth, and which will make it possible to make firm conclusions about the impacts of *doi moi* on economic development, its relationship with inequality and the patterns of growth.

References

Abreu, M., de Groot, H. and Florax, R. (2004). Growth and Space, Discussion Paper: 2004-129/3, Tinbergen Institute.

Aday, L. A. and Anderson, R. (1975). Development of Indices of Access to Medical Care, *Health Administrative Press*, Ann Arbor.

Ahrend, R. (2002). Speed of Reform, Initial Conditions, Political Orientation, or What?, Explaining Russian Regions' Economic Performance, DELTA Working Paper No. 2002-10, Départment et labaratoire d'économie théorique et appliqué, Paris.

Akita, T. (2003). Decomposing Regional Income Inequality in China and Indonesia using Two-Stage Nested Theil Decomposition Method, *Annals of Regional Science*, 37 (1), pp.55-77.

Alesina, A. and D. Rodri. (1994), Redistributive Politics and Economic Growth, *Quarterly Journal of Economics*, Vol. 436, pp. 465-490.

Amos, O. M. JR. (1988). Unbalanced Regional Growth and Regional Income Inequality in the Latter Stages of Development, *Regional Science and Urban Economics*, Vol. 18, pp. 554-566.

Anand, S. (1983). Inequality and Poverty in Malaysia: Measurement and Decomposition, a World Bank Research Publication. Oxford University Press, New York.

Anselin L. and A. K. Bera, R. J. G. M. Florax, and M. Yoon (1996), Simple Diagnostic Tests for Spatial Dependence, *Regional Science and Urban Economics*, Vol. 26, pp. 77-104.

Anselin, L. (1988a). *Spatial Econometrics: Methods and Models,* Kluwer Academic Publisher, Dordrecht.

Anselin, L. (1988b), Model Validation in Spatial Econometrics: A Review and Evaluation of Alternative Approaches, *International Regional Science Review,* Vol. 11, pp. 279-316.

Anselin, L. (1991). The Performance of Tests for Spatial Dependence in a Linear Regression, Report 91-13, National centre for Geographic Information and Analysis, Department of Geography, Uni. of California at Santa Barbara, CA 93106.

Anselin, L. (1999). Rao's Score Tests in Spatial Econometrics*, Journal of Statistical Planning and Inference.*

Anselin, L. (2001). Spatial Econometrics, in Baltagi B. (ed.), *Companion to Econometrics*, Basil Blackwell, Oxford.

Anselin, L. and A. K. Bera (1998). *Spatial Dependence in Linear Regression Models with an Application to Spatial Econometrics*, in Ullah A., Giles D.E.A. (eds.), Handbook of Applied Economic Statistics, Springer, Berlin.

Anselin, L. and R. Florax (1995). Small Sample Properties of Tests for Spatial Dependence in Regression Models: Some Further Results, in: Anselin, L, Florax, JGM (eds.) *New Directions in Spatial Econometrics. Berlin*, New York, pp. 21-71.

Anselin, L. and S. J. Rey (1997). Introduction to the Special Issue on Spatial Econometrics, *International Regional Science Review*, Vol. 20, pp. 1-7.

Arbia, G. and Piras, G. (2004). Convergence in Per Capita GDP across European Regions Using Panel Data Models Extended to Spatial Autocorrelation Effect, Paper presented at the European Regional Science Association (ERSA) Conference, Porto.

Armstrong, H. (1995). An Appraisal of the Evidence from Cross-sectional Analysis of the Regional Growth Process within the European Union, in Armstrong H., Vickerman R. (Eds.), *Convergence and Divergence among European Union*, Pion, Londres.

Atkinson, A. B. (1996). Bringing Income Distribution in from the Cold, The presidential address to the Royal Economic Society, Swansea, April.

Azzoni, C. R., N. Menezes-Filho, T. A. de Menezes, and R. Silveira-Neto (2000). Geography and Income Convergence among Brazilian States, Inter-American Development Bank, Research Network Working Paper # R-395.

Baltagi, B. H. and D. Li, (1999). Prediction in the Panel Data Model with Spatial Correlation, in L. Anselin and R. J. G. M Florax (eds.), *New Advances in Spatial Econometrics*.

Baltagi, B. H. (2001). *Econometrics Analysis of Panel Data,* Willey, Chichester.

Baltagi, B. H. and D. Li, (2001a). LM Test for Functional Form and Spatial Error Model*, International Regional Science Review*, Vol. 24(2), pp. 194-225.

Baltagi, B. H. and D. Li, (2001b). Double Length Artificial Regressions for Testing Spatial Dependence, *Econometric Review*, Vol. 20(1), pp. 31-40.

Banerjee, A. and E. Duflo, (1999). *Inequality and growth: What Can the Data Say?* Mimeo, MIT.

Barro R. J. and X. Sala-I-Martin, (1991). Convergence across States and Regions, *Brooking Papers on Economic Activity*, Vol. 1, pp. 107-182.

Barro R. J. and X. Sala-I-Martin, (1992). Convergence, *Journal of Political Economy*, Vol. 100, pp. 223-251.

Barro R. J. and X. Sala-I-Martin, (1995). *Economic Growth*, McGraw-Hill, New York.

Barro, R. J. (1991). Economic growth in a cross section of countries, *The Quarterly Journal of Economics,* Vol. 106(2), 407-443.

Barro, R. J. (2000), Inequality, Growth in a Panel of Countries, *Journal of Economic Growth*, Vol. 5, pp. 5-32.

Barro, R. J. and Lee, (2000). *World Development Indicators*, World Bank Poverty Net.

Barro, R. J. and X. Sala-i-Martin, (1998). *Economic Growth,* McGraw-Hill, New York.

Bell K. P. and N. R. Bockstael, (2000). Applying the Generalized – Moments Estimation Approach to Spatial Problems Involving Micro level Data, *Review of Economics and Statistics*, Vol. 82, pp. 72-82.

Bell, K.P. and N. Bockstael, (2000). Applying the Generalized Methods of Moments Approach to Spatial Problems Involving Micro-level Data, *Review of Economics and Statistics*, Vol. 821, pp. 72-82.

Berkowits, D. and D. N. Dejong, (2001). The Evolution of Market Integration in Russia*, Economics of Transition*, Vol. 9(1); pp. 87-104.

Bernard, A. and Jones, C. (1996). Productivity and Convergence across U.S. States and Industries, *Empirical Economics*, Vol. 21, pp. 113-135.

Berra, A. K. and M. J. Yoon, (1993). Specification Testing with Misspecified Alternatives*, Econometric Theory*, Vol. 9, pp. 649-658.

Binh, T. Nguyen, J. W. Albrecht, S. B. Vroman and M. D. Westbrook, (2003). A Quantize Regression Decomposition of Urban-Rural Inequality in Vietnam, Research paper, Department of Economics, Georgetown University.

Bourguignon, F., (1979). Decomposable Income Inequality Measures, *Econometrica*, Vol. 47(4), pp. 901–920.

Bradshaw, M. J. and K. Vartapetov, (2003). A New Perspective on Regional Inequalities in Russia*, Eurasian Geography and Economics*, Vol. 44 (6), pp. 403-429.

Brooks, C. H. (1973). Associations among Distance, Patient Satisfaction, and Utilization of Two Types of Inner-City- Clinics, *Medical Care*, Vol. 11, pp. 373-383.

Burridge, P., (1981). Testing for a Common Factor in a Spatial Autoregressive Model, *Environment and Planning Analysis*, Vol. 13, pp 795-800.

Cai, F., D. Wang and D. Yang, (2002). Regional Disparity and Economic Growth in China: The Impact of Labour Market Distortions, *China Economic Review*, Vol. 13 (2), pp. 197-212.

Canadian Institute for Health Information, (1998). National Consensus Conference on Population Health Indicators, Final Report, Ottawa, Ontario.

Carlino, G. and Mills, L. (1996a). Convergence and the U.S. States: A Time-Series Analysis, *Journal of Regional Science*, Vol. 36, pp. 597-616.

Carlino, G. and Mills, L. (1996b). Testing Neoclassical Convergence in Regional Incomes and Earnings, *Regional Science and Urban Economics*, Vol. 26, pp. 565-590.

Case, A. C., (1991). Spatial Patterns in Household Demand*, Econometrica*, Vol. 59(4), pp. 953-965.

Case, A. C., J. Hines and H. Rosen, (1993). Budget Spillovers and Fiscal Policy Independence: Evidence from the States, *Journal of Public Economics*, Vol. 52, pp. 285-307.

Chen J. and B. M. Fleisher, (1996). Regional Income Inequality and Economic Growth in China*, Journal of Comparative Economics*, Vol. 22 (2), pp. 141-164.

Chen, M. K., (1973). Access to Health Care: A Preliminary Model, Unpublished Manuscript, National Center for Health Services Research, Hyattsville, Maryland.

Chen, M. K., (1978). A Quantitative Index of Access to Primary Medical Care for Health Planning, *Socio – Economic Planning Sciences*, Vol. 12, pp. 259-301.

Cliff, A and K. Ord, (1972). Testing for Spatial Autocorrelation among Regression Residuals*, Geographical Analysis*, Vol. 4, pp. 267-284.

Cliff, A. and J. K. Ord, (1973). *Spatial Autocorrelation,* London: Pion.

Congress of the Communist Party of Vietnam (1976), Official Documents of the Congress, National Political Publisher, Hanoi.

Conley, T. G., (1999). GMM Estimation with Cross Sectional Dependence, *Journal of Econometrics*, Vol. 92, pp. 1-45.

Currie, Janet, and Duncan Thomas, (1995). Does Head Start Make a Difference? *American Economic Review*, Vol. 85(3), pp. 341-64.

Currie, Janet, and Duncan Thomas, (2000). School Quality and Longer-Term Effects of Head Start, *Journal of Human Resources*, Vol. 35(4), pp. 755-74.

Dacy, D., (1986). *Foreign Aid, War and Economic Development: South Vietnam, 1955-1975*, Cambridge, Cambridge University Press.

Démurger, S., J. D. Sachs and W. T. Woo, (2002). Geography, Economic Policy, and Regional Development in China, NBER Working Paper No. 8897, National Bureau of economic Research, Cambridge.

Do, Duc Dinh, (1995). The Social Impact of Economic Reconstruction in Vietnam: A Selected Review, Labor Institutions and Development Programme, Hanoi: DP/81/1995, ISBN 92-9014-591-9.

Dolinskaya, I., (2002). Transition and Regional Inequality in Russia: Reorganization or Procrastination? IMF Working Paper WP/02/169, International Monetary Fund, Washington D.C.

Dollar, D., P. Glewwe and J. Litvack (Eds.), (1998). Household Welfare and Vietnam's Transition, World Bank, Regional and Sectoral Studies.

Driscoll, J. and A. Kraay, (1998). Consistent Covariance matrix Estimation with Spatially Dependent Panel Data, *Review of Economics and Statistics*, Vol. 80, pp. 549-560

Ducan, R. and X. Tian, (1999). China's Inter-provincial Disparities, *Communist and Post-communist Studies,* Vol. 32 (2), pp. 211-224.

Easterlin, R., (1960a). International Differences in Per Capita Income, Population and Total Income, 1840-1950, in National Bureau of Economic Research Trends in the American Economy in the Nineteenth Century, Princeton University Press, Princeton.

Easterlin, R., (1960b). Regional Growth of Income, in Kuznets, S., Miller, A., and Easterlin, R. (Eds.) *Population Redistribution and Economic Growth in the United States, 1870-1950*. The American Philosophy Society, Philadelphia.

Easterly, W., (2001). The Middle Class Consensus and Economic Development, *Journal of Economic Growth*, Vol. 6, pp. 317-335.

Epprecht M. and A. Heinimann eds., (2004). Socioeconomic Atlas of Vietnam: A Depiction of the 1999 Population and Housing Census, Swiss National Centre of Competence in Research (NCCR) North-South, University of Berne, Berne.

Ertur C. and W. Koch, (2006). Regional Disparities in the European Union and the Enlargement Process: An Exploratory Spatial Data Analysis, 1995-2000, Forthcoming in *Annals of Regional Science*, Doi: 10.1007/s00168-006-0062-x.

Ertur C., J. Le Gallo and C. Baumont, (2006). The European Regional Convergence Process, 1980-1995: Do Spatial Regimes and Spatial Dependence Matter? *International Regional Science Review*, Vol. 29. pp. 3-34.

European Commission, (1997). Directorate General for Employment, Industrial Relations and Social Affairs, Public Health in Europe, Luxembourg: Office for Official Publications of the European Commissions.

Evans, P., (1998). Using Panel Data to Evaluate Growth Theory, *International Economic Review*, Vol. 39(2), pages 295-306.

Evans, P. and G. Karras, (1996). Convergence Revisited, *Journal of Monetary Economics*, Vol. 37, No. 2, pp. 249-265.

Fedorov, L., (2002). Regional Inequality and Regional Polarization in Russia, 1990-99, *World Development*, Vol. 30 (3), pp. 443-456.

Fforde, A. and S. de Vylder, (1996). From Plan to Market: The Economic Transition in Vietnam, Boulder, The Westview of the Press.

Fingleton, B. (ed.), (2003). *European Regional Growth*, Berlin, Heidelberg and New York, Springer.

Fisher, F. M., (1956). Income Distribution, Value Judgements and Welfare, *Quarterly Journal of Economics*, Vol. 70, pp. 380 – 424.

Florax, R. and P. Nijkamp, (2004). Misspecification in Linear Spatial Regression Models, in Kempf-Leonard K. (ed.), *Encyclopedia of Social Measurement*, Academic Press, San Diego.

Florax, R. J. G. M, H. Folmer and S. J. Rey, (2002). Specification Searches in Spatial Econometrics: the Relevance of Hendry's Methodology, Urban/Regional 0202001, EconWPA.

Flury, B., (1988). *Common Principal Components and Related Multivariate Models*, John Wiley, New York.

Folmer, H., (1986). *Regional Economic Policy: Measurement of Its Effects*, Dordrecht: Martinus Nijkoff.

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

Friedman, M. (1955). The Role of Government in Education, in *Economics and the Public of Interests*, Robert A. Solo Ed., Rutgers University Press.

Gabriele, A., (2006). Social Services Policies in a developing Market Economy Oriented towards Socialism: The Case of Health System Reforms in Vietnam, *Review of International Political Economy*, Vol. 13; pp. 258-289.

Galor, O. and J. Zeira, (1993). Income Distribution and Macroeconomics, *the Review of Economic Studies*, Vol. 60, pp. 33-52.

General Statistical Office of Vietnam (GSO), (2000). *Statistical Year Books*, Statistical Publishing House, Hanoi, Vietnam.

General Statistical Office of Vietnam (GSO), (2001). *Statistical Year Book 2001,* Statistics Publishing House, Hanoi, Vietnam.

General Statistical Office of Vietnam (GSO), (2002). *Statistical Year Book 2002,* Statistics Publishing House, Hanoi, Vietnam.

General Statistical Office of Vietnam (GSO), (2003). *Statistical Year Book 2003,* Statistics Publishing House, Hanoi, Vietnam.

General Statistical Office of Vietnam (GSO), (2004). *Statistical Year Book 2004,* Statistics Publishing House, Hanoi, Vietnam.

General Statistical Office of Vietnam (GSO), (2005). *Statistical Year Book 2005,* Statistics Publishing House, Hanoi, Vietnam.

General Statistical Office of Vietnam (GSO), (2006). *Statistical Year Book 2006,* Statistics Publishing House, Hanoi, Vietnam.

Government of Vietnam-Donor-NGO Poverty Working Group (PWG) (1999). *Vietnam: Attacking Poverty*, Hanoi, World Bank.

GSO (General Statistical Office) (Various years from 1990 to 1999). *Statistical Yearbook*, Statistical Publishing House, Hanoi.

GSO (General Statistical Office) and UNICEF (United Nations Children's Fund) (2004). *Data on Social Statistics in the Early Years of the 21st Century*, Statistical Publishing House, Hanoi.

Heltberg, R., (2003). Spatial Inequality in Vietnam: A Regression-Based Decomposition, the World Bank Research Paper, the World Bank.

Holtz-Eakin, D. (1994). Public Sector Capital and the Productivity Puzzle*, Review of Economics and Statistics*, Vol. 76, pp. 12-21.

Hordijk, L. (1974). Spatial Correlation in the Disturbance of a linear interregional model, *Regional Science and Urban Economics*, Vol. 4, pp. 117-140.

Im, K-S., H. Pesaran and Y. Shin, (2003). Testing for Unit Roots in Heterogeneous Panels, *Journal of Econometrics*, Vol. 115, pp. 53-74.

Justino, P., J. Litchfield and Y. Niimi, (2004). Multidimensional Inequality: An Application to Brazil, PRUS Working Paper No. 24, Department of Economics, University of Sussex.

Kaldor, N., (1957). *A* Model of Economic Growth, *Economic Journal*, Vol. 57, pp. 591-642.

Kelejian H. H. and D. P. Robinson, (1992). Spatial Autocorrelation: A New Computationally Simple Test with an Application to Per Capita County Police Expenditures, *Regional Science and Urban Economics*, Vol. 22, pp. 317-333.

Khan, A. A., (1985). Analyzing Spatial Disparities in Access to Health Care: A Methodology with Application in Bangladesh, *Geo Journal*, Vol. 10(1), pp. 97-107.

Kolm, S. Ch., (1977). Multidimensional Egalitarianism, *Quarterly Journal of Economics*, Vol. 91, pp.1-13.

Krugman P., (1991). Increasing Returns and Economic Geography, *Journal of Political Economy*, Vol. 99, pp. 483-499.

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

Le Gallo, J. and C. Ertur, (2003). Exploratory Spatial Analysis of the Distribution of Regional Per Capita GDP in Europe, 1980-1995, *Papers in Regional Science*, Vol. 82, pp. 175-201.

LeSaga, J. P., (1999). *Spatial Econometrics*, Department of Economics, University of Toledo Publisher.

Levin, A., C-F Lin and C. Chu, (2002). Unit Root Test in Panel Data: Asymptotic and Finite Sample Results, *Journal of Econometrics*, Vol. 108, pp. 1-24.

Li, H. and H. Zou, (1998). Income Inequality is Not Harmful for growth: Theory and Evidence, *Review of Development Economics*, Vol. 2, pp.318-334.

Liu, Amy Y. C., (2001). Market, Inequality and Poverty in Vietnam, *Asian Economic Journal*, Vol. 15, (2), 217-235.

Ljunggen, B., (ed.), (1993). *The Challenge of Reform in Indochina*, Cambridge, Mass: Harvard University Press.

Lope'z-Bazo E., E. Vaya, A. J. Mora and J. Suriach, (1999). Regional Economic Dynamic and Convergence in the European Union, *Annals of Regional Science*, Vol. 33, pp. 343-370.

Lugo. M. A., (2004). On Multivariate Distribution of Well-Being: The Case of the Argentine Provinces in the 1990s, Paper Represented at the General Conference of the International Association for Research in Income and Wealth, Cork, Ireland, August 2004.

Lydall, H. F., and J. B. Lansing, (1959). A Comparison of the Distribution of Personal Income and Wealth in the United States and Great Britain, *American Economic Review*, Vol. 49, pp.43-67.

Maasoumi, E., (1986). The Measurement and Decomposition of Multidimensional Inequality, *Econometrica*, Vol. 54(4), pp. 991-997.

Maasoumi, E., (1989). Composite Indices of Income and Other Development Indicators: A General Approach, *Research on Economic Inequality*, Vol. 1, pp. 269-288.

Maasoumi, E. and G. Nickelsburg, (1988). Multivariate Measures of Well-being and an Analysis of inequality in the Michigan Data, *Journal of Business and Economic Statistics*, American Statistical Association, Vol. 6(3), pp. 326-334.

Maasoumi, E. and J-H. Jeong, (1985). The Trend and the Measurement of Inequality over Extended Periods of Accounting, *Economic Letters*, Vol. 19, pp. 295-301.

Mackenbach, J. P. and A. E. Kunst, (1997). Measuring the Magnitude of Socio-Economic Inequality in Health: An Overview of Available Measures Illustrated With Two Examples from Europe, *Social Sciences & Medicine*, Vol. 44, No. 6, pp. 757-771.

Magrini, S., (2004). *Regional Convergence*, in Handbook of Regional and Urban Economics, Henderson V., Thisse J.F.(Eds.), Elservier: New York.

McCabe, G. P., (1984). Principal Variables, *Technometrics,* Vol. 26, pp.137-144

Miller, J. R. and I. Genk, (2005). Alternative Regional Specification and Convergence on U.S. Regional Growth Rate, *Annals of Regional Science*, Vol. 39, pp. 241-252.

Minot, N., B. Baulch and M. Epprecht, (2003). Poverty and Inequality in Vietnam: Spatial Patterns and Geographic Determinants, Research paper, International Food Policy Research Institute, Washington, D.C. and Institute of Development Studies, University of Sussex, Brighton, 87p.

Morris, M. D., (1979). Measuring the Condition of the World's Poor: Physical Quantity of Life Index, Overseas Development Council, Washington, D.C.

Morrison, Donald F., (1967). *Multivariate Statistical Method*, McGraw-Hill, New York.

National Centre of Social Sciences and Humanities of Vietnam (2001), doi moi and Human Development in Vietnam, Hanoi.

Nghia, L.X., Van Donge and Litvack, J. (1999). Fostering High Growth in a Low Income Country: Program Aid to Vietnam, Stockholm: Sida.

Nguyen Thanh Son, (2002). Education in Vietnam in the period 1996-2010: A tool for poverty eradication and social development, International Conference: Developing Skill for the New Economy, 17-19 October 2002, Winnipeg-Manitoba, Canada.

Nijkamp, P., (1978). Quality of Life and Changes in Human Settlement Patterns, in Hansen, N. M. Human Settlement System: International Perspective on Structure, Change and Public Policy. Cambridge (Mass.): Ballinger.

Nijkamp, P., (1988). Multidimensional Spatial Data and Decision Analysis, Ann Arbor, Mich: University Microfilms International.

Ord, J. K. (1975), Estimation Methods for Models of Spatial Interaction, *Journal of the American Statistical Association*, Vol. 70, pp. 120-126.

Panizza, U., (2002), Income Inequality and Economic Growth: Evidence from American Data*, Journal of Economic Growth*, Vol. 7, Issue 25, pp.25-41.

Pedroni, P. and J. Y. Yao, (2005). Regional Income Divergence in China, Working papers, Baker Institute and department of Economics, Rice University, 6100 Main St., Houston.

Perotti, R., (1996). Growth, Income Distribution and Democracy: What the Data Say? *Journal of Economic Growth*, Vol. 1, pp. 149-187.

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

Pham Manh Hung, Truong Viet Dzung, Dahlgren, G. and Tran Tuan, (2001). Vietnam: Efficient, Equity Oriented Financial strategy for Health, in T. Evans, M. Withead and M. Pradhan (eds.) (2002) "How much does it cost to relieve the poor from private contributions from basic education and health?" Unpublished manuscript, World Bank, Hanoi.

Phillips, P., (1994). CPC-Common Principal Component Analysis Program, Biology Department, University of Texas at Arlington, Box. 19498.

Phillips, P., (1997). *Common Principal Analysis*, University of Texas Arlington.

Phong, D. and M. Beresford ,(2000). *Economic Transition in Vietnam: Trade and Aid in the Demise of a Centrally Planned Economy*, London: Edward Elgar (Table 2.10).

Projector, D. S., and G. S. Weiss, (1966). Survey of Financial Characteristics of Consumer, Unpublished Report, Board of Governors of the Federal Reserve System, Washington DC.

Quadrado, L, Henk Folmer and Loman S, (2001). Regional Inequality in the Provision of Healthcare in Spain, *Environment and Planning*, Vol. 33: pp. 783-793

Quadrado, L. (1999). *Welfare inequality, regionalization and welfare policy: Measurement and analysis for Spain,* PhD thesis, Wageningen University, The Netherlands

Quadrado, L., S. Loman and H. Folmer, (2001). Multi-dimensional analysis of regional inequality: The case of higher educational facilities in Spain, *Regional Science*, Vol. 80, pp. 189-209.

Quah, D., (1996). Regional Convergence Clusters across Europe, *European Economic Review*, Vol. 40, pp. 951-995.

Quah, D., (2002). One Third of the World's Growth and Inequality, Economics Department, CEPR Discussion Paper 2002:3316.

Ram, R., (1982). Composite Indices of Physical Quality of Life, Basic Needs Fulfillments and Income: A Principal Component Representation, *Journal of Development Economics*, Vol. 11. pp. 227-247.

Ram, R. (1982). Composite Indices of Physical Quality of Life, Basic Needs Fulfillment and Income: A Principal Component Representation, *Journal of Development Economics*, Vol. 11, pp. 227-47.

Rey, S. J., B. D. Montouri, (1999). U.S. Regional Income Convergence: A Spatial Econometric Perspective, *Regional Studies*, Vol. 33, pp. 145-156.

Riddell, W. C., (2004). The Social Benefits of Education: New Evidence on an Old Question, Department of Economics, University of British Columbia (Paper presented at the conference "Taking Public University Seriously", University of Toronto, December 3-4, 2004).

Schneider, M. C., C. C. Salgado, J. Bacallao, E. Loyola, et al., (2004). Methods for Measuring Health Inequalities (Part I), *Epidemiological Bulletin*, Vol. 25(4).

Segall, M., Tipping, G., Lucas, H., Truong Viet Dung, Nguyen Thanh Tam, Dao Xuan Vinh and Dao Lan Huong, (2000). Healthcare Seeking by the Poor in Transitional Economies: The Case of Vietnam, Brighton, UK: Institute of Development Studies.

Sen, A. K., (2001). Why Health Equity? Keynote address to the Third International Conference on "The Economics of Health: Within and Beyond Health Care", York, UK, 23 July.

Sen, A. K., (1980). Equality of What?, in McMurrin S: Tanner Lectures on Human Values. Cambridge: Cambridge University Press.

Sen, A. K., (1985a). A Sociological Approach to the Measurement of Poverty: A Reply to Professor Peter Townsend, Oxford Economic Paper, 37.

Sen, A. K., (1985b). *Commodities and Capabilities*, North-Holland, Oxford U. Press

Sen, A. K., (1987a). *Commodities and Capabilities*, New Delhi: Oxford University Press.

Sen, A. K., (1987b). From Income Inequality to Economic Inequality*, Southern Economic Journal*, Vol. 64.

Sen, A. K., (1988). Freedom of Choice: Concepts and Content, *European Economic Review*, Vol. 32; 269-294.

Sen, A. K., (1992). *Inequality Re-examined,* Oxford: Clarendon Press, New York: Russell Sage Foundation, and Cambridge. MA: Harvard University Press.

Sen, A. K., (1993). Well being and Capability, in Nussbaum M & Sen A K: *The Quality of Life*. Oxford: Oxford University Press.

Sen, A. K., (1997). From Income Inequality to Economic Inequality, *Southern Economic Journal,* Vol. 64, pp. 384-401.

Sen, A. K., (1998). Should Inequality and Poverty Measures be Decomposable*,* A Report on the Second Kumar Chakravarti Memorial Lecture, Calcutta Statistical Association Bulletin, Vol. 48.

Sepehri, A. and A H. Akram-Lodhi, (2002). A Crouching Tiger? A Hidden Dragon? Transition, Saving and Growth in Vietnam, 1975-2006, Working Paper 359, the University of Manitoba, Winnipeg, Canada R3T 2N2.

Shannon, G. W., (1981). Accessibility to Medical Care among Urban Native Americans in a Large Metropolitan Areas, Research in Contemporary and Applied Geographic Discussion Series No. 5 (1). Department of Geography, State University of New York, Bin-Hampton, NY.

Shirakawa and Morimoto, (1991). Lifestyle effect on total IgE - Lifestyles have a cumulative impact on controlling total IgE levels*, Allergy*, Vol. 46(8), pp. 561-569.

Shorrocks, A. F., (1980), The class of Additively Decomposable Inequality Measures, *Econometrica*, Vol. 48(3), pp.613–625.

Solanko, L., (2003). An Empirical Note on Growth and Convergence across Russian Regions*,* BOFIT Discussion Papers No. 9, Institute for Economies in Transition, Bank of Finland, Helsinki.

Taussig, M. K., (1973). Alternative Measures of the Distribution of Economic Welfare*,* Princeton University, Department of Economics.

Tian, X., (1999). Market Orientation and Regional Economic Disparities in China*, Post Communist Economies*, Vol. 11 (2), pp. 161-172.

United Nations, (2003). Health Care Financing for Vietnam, United Nations Country Team Vietnam, June 2003, Hanoi.

United Nations in Vietnam, (2003). *Millennium Development Goals: Closing the Millennium Gaps*, Hanoi November 2003.

Vietnam Academy of Social Sciences, (2006). *National Human Development Report*, National Political Publishing House, Hanoi.

Vietnam Academy of Social Sciences, (2006). Vietnam's *Human Development Report*, Hanoi.

Vietnam Government, (2005). *Vietnam Achieving the Millennium Development Goals*, Socialist Republic of Vietnam.

Weeks, M. and J. Y. Yao, (2003). Provincial Conditional Income Convergence in China, 1953-1997: A Panel Data Approach*, Econometric Review*, Vol. 22 (1), pp. 59-77.

WHO (1978), *Financing of Health Services*, World Health Organization, Geneva

Wiley, J. A. and T. C. Camacho ,(1980). Life-style and Future Health: Evidence from the Alameda County Study*, Preventive Medicine*, Vol. 9, pp. 1-21.

Williamson, J., (1965). Regional Inequality and the Process of National Development, *Economic Development and Cultural Change*, Vol. 4, pp.3-47.

World Bank, (1990). Vietnam: *Vietnam Stabilization and Structural Reforms*, Washington, the World Bank.

World Bank, (1995). *Vietnam Development Report 1995*, World Bank, Vietnam.

World Bank, (2000). *Vietnam Development Report* 2000, World Bank, Vietnam.

World Bank, (2001a). Vietnam Economic Monitor Spring 2001, hppt://www.worldbank.org.vn/rep30.

World Bank, (2002). *Vietnam Development Report* 1995, World Bank, Vietnam.

World Bank, (2004). Vietnam, *Vietnam Development Report*, Washington, the World Bank.

World Bank, (2005). *Vietnam Development Report-Poverty -2004*, World Bank, Vietnam.

World Bank, (2006). *World Development Report 2006*, World Bank, Washington DC.

World Bank, (2006). *World Development Report 2006: Equity and Development*, World Bank and Oxford University Press, New York 10016.

World Bank. *World Development Report*, various years in 1990s and 2000s.

Xu, L. C. and H. F. Zou, (2000). Explaining the Changes of Income Distribution in China*,* *China Economic Review*, Vol. 11, pp. 149-170.

Zandvakili, S., (1992). Generalised Entropy Measures of Long-run Inequality and Stability among Male Headed Households*, Empirical Economics*, Vol. 17, pp. 565-581.

Zandvakili, S., (1999). Income Inequality among Female Heads of Households: Racial Inequality Reconsidered*, Econometrica*, Vol. 66, pp. 119-133.

Zhai, F., T. Hertel and Z. Wang, (2003). Labour Market Distortions, Rural-Urban Inequality and the Opening of China's Economy*, Journal of Development Studies*, Vol. 37(3), pp. 85-98.

Zhang, X. and R. Kanbur, (2001). What Difference Do Polarization Measures Make? An Application to China*, Journal of Development Studies*, Vol. 37 (3), pp. 85-98.

Zhang, Z., A. Liu and S. Yao, (2001). Convergence of China's Regional Income: 1952-1997, *China Economic Review*, Vol. 12, pp. 243-258.

Summary

This study had multiple objectives. First, it aimed at examining regional inequality in several welfare variables, such as health and education during the 1998-2004 period, and per capita income over time from 1990 to 2006. The first objective of the study was to gain a better understanding about the magnitude and evolution of regional inequality in health, education and per capita income. In addition, the study paid attention to the levels and trends of between- and within-region differences in these welfare variables separately. The second objective of the study was to explore the relationship between regional income inequality and regional growth during the 1990-2006 period in order to provide knowledge about empirical evidence of the relation between regional income inequality and regional economic growth in Vietnam. The third and final objective of the study was to examine the regional growth patterns in the country for the growth episode stretching from 1990 to 2006. Also in the third objective of the study, the role of space and spatial autocorrelations were considered in order to provide a better understanding about the spatial interaction between regions in Vietnam and its impact on regional economic growth. To attain these objectives, the study employed various methodological methods in order to measure and analyze regional inequality in welfare variables, to explore the empirical relationship between inequality and growth, and to test for growth patterns in the Vietnam economy. The study was subdivided into eight chapters. Besides the introduction provided in chapter one, an overview, conclusions, discussion and recommendations given in chapter eight, the other six chapters, from chapter two to chapter seven, covered all issues relating to characteristics of reform and welfare in Vietnam, the literature, methodological approach and empirical analysis.

Economic reforms and changes in welfare in Vietnam were elaborated and discussed in chapter two. The main focus of this chapter was to discuss the macro-economic situation, as well as the implementation of the economic reforms called *doi moi*, and its impact on people's welfare. As discussed in the chapter, before the *doi moi*, the welfare system in Vietnam functioned quite well. Most Vietnamese people had equal access to most welfare services provided by the government, irrespective of religion, ethnicity or social status. Similar to other communist countries, welfare distribution in Vietnam before the *doi moi* was relatively equal, although the quality of the welfare system with respect to services offered was still questionable in terms of catching up to the international standard. At the time, most Vietnamese people enjoyed free access to the healthcare system. Pupils and students got free education at all levels. As a result, the free systems culminated in good health records attained throughout the country, high enrolment rates and a high rate of literacy among the Vietnamese people compared with other developing countries with the same or slightly higher level of development.

The situation and condition of welfare, such as education and health after the *doi moi,* were also discussed and compared with the condition of welfare before *doi moi* to evaluate the impact of high and relatively high rates of economic growth of 6-9% per annum on a changing welfare system, by explanatorily analyzing the welfare system, using Gini coefficients. This showed that regional inequality in the country had generally changed.

The methodological issue regarding measuring and analyzing regional welfare inequality was dealt with in chapter three. Apart from discussing the theoretical approach

and empirical measurement models, the chapter also provided a short review of the literature on multidimensional regional inequality. From the various approaches to welfare inequality measurement and analysis, the present study chose the multidimensional approach, using Theil's second index as the inequality measure. The rationale for using Theil's second index (measure) and the valuable properties of the Theil's second index were also discussed in this chapter. Theil's second measure satisfies the property of additive decomposability and is very informative about the status of the distribution of well-being. Additive decomposability property meant that overall inequality could be decomposed into within- and between-region inequality, which would be very meaningful for understanding magnitudes and trends of regional inequality. Theil's second measure also satisfied several desirable properties as measuring regional income/welfare inequality, such as mean independence, population-size independence, and the Pigou-Dalton principle of transfers. In order to compute Theil's second measure of inequality, there is a need for estimating a composite index, based on the aggregate function by Maasoumi, in which the indicators' coefficients and their associated weights are to be estimated. To do so, the chapter also developed the measurement models based on the principal component models (PCA) and partial common principal component models (PCPC), developed and used by Flury. These two types of models were discussed in detail in order to help the readers to comprehend the subsequent empirical parts of the thesis.

The measurement and analysis of regional inequality for two welfare variables, namely health and education, were given in chapter four and chapter five, respectively.

In chapter four, we proceeded with the measurement and analysis for health facilities and health status, and also examined spillover effects, for example, that health facilities in one region would affect people's health status in the first-order adjacent province(s). We carried out several tests for spatial dependence among these two sub-components of health. The results in chapter four revealed that total regional inequality had different patterns for healthcare facilities and health status during the 1998-2004 period. Generally, regional inequality in both the healthcare facilities and health status moderately decreased in Vietnam over the 1998-2004 period. However, there was a sign of increase during this stretch of time, as the results found in chapter four showed a slight increase in inequality in both components of health over the 1998-2002 period, and then it decreased in the next period.

Empirical findings in chapter four also proved the importance of spillover effects prevailing in the case of the health component. Healthcare services seemed to be more equally distributed in the case of contiguity than in the case of non-contiguity, as Theil's second measure in the former was lower than that in the latter. This finding indicated that if we ignored the existence of spillover effects, the regional inequality would be exaggerated.

The tests for spatial dependence among health components were also conducted in chapter four for several specifications regarding the relationship between health status and health facilities. The results of the tests confirmed the existence of spatial effects between the two dimensions of health. The test also proved that the spatial error model was an appropriate specification for analysis of the relationship between the composite index with respect to health status and the composite index with respect to health facilities, while the spatial lagged model was a more appropriate specification to analyze the relationship between the composite index with respect to health status and the indicators of health facilities.

Following chapter four, the measurement and analysis of multidimensional regional inequality for the education component were carried out in chapter five. However, different from chapter four, contiguity effects were not considered in this chapter, because we argued that due to mandatory education applied in Vietnam, pupils were not allowed to attend school in other regions (provinces) than they were registered to live (even though in practice pupils would seek education in other regions other than the one they lived, but this phenomenon was not common, so we did not have statistics for analysis). Thus, as a result, we did not take spatial dependence into consideration.

For regional inequality in education facilities, the research showed that total inequality, its between-region term and within-region components all go into opposite directions between the two periods. However, the change in the within-region inequality was in contrast with the changes in total regional inequality. The between-region component as the total regional inequality and the between-region component in education facilities first decreased during the period 1998-2002 and then increased during the period 2002-2004, while the within-region inequality increased during the 1998-2002 period, then decreased during the 2002-2004 period. As research findings revealed, the distribution of education facilities among the regions in Vietnam was a bit more unequal over time. Meanwhile, the change over time of the within-region component of inequality was very moderate.

For regional inequality in enrolments, our findings showed that the pattern of total regional inequality and its decomposed terms were the same as all of them increased in the first period (1998-2002), and declined in the next period (2002-2004). The structural shares of between-region and within-region inequality in the total inequality told us that in all three years under study the between-region term made up a larger part in the total inequality. These implied that the differences in enrolments between regions were more significant to the total inequality than the differences in enrolments within regions.

The comparative composite index with respect to education facilities and the composite index with respect to enrolments presented in chapter five revealed that there was a sharp contrast between the two dimensions of education, as the most-favored regions with respect to enrolments were those in the lowland areas with least and medium-favored education facilities. The least-favored regions with respect to the education enrolments were those in the upland areas, which were very high-favored with respect to education facilities in terms of the composite index.

Chapter six presented the measurement and analysis of the regional income inequality, and explored the evolution of income disparities, its decomposed terms as well as its relationship with economic growth among regions in Vietnam in the period from 1990 to 2006. The findings confirmed that generally the regional income inequality in the country moderately increased over time. The decomposition of the total inequality revealed that within-region inequality decreased, while between-region increased over time. However, the size of within-region income inequality differed among regions in the country, as within-region Theil's second inequality measure showed different trends.

The empirical exploration of the relationship between regional inequality and regional growth found revealed different patterns of the relationship. The estimates from the panel data regression have found a positive relationship between the rate of economic growth and the initial level of per capita income in all the cases studied. Meanwhile, using the pooled OLS technique, we found negative relationships between growth and inequality for all the cases studied: without both controls and regional dummies, with the controls and no regional dummies, and with both controls and regional dummies.

The test for the validity of the method (either the fixed effects or random effects is more appropriate and preferred) revealed that the fixed effects method was appropriate and preferred for the cases: without both the controls and dummies and with controls but no dummies in the growth model, while the random effects technique was more appropriate and more preferred for the case with both the controls, regional and time dummies. Interestingly, estimates by both the fixed effects and random effects methods revealed that there was both a positive and negative relationship between regional growth rate and initial level of per capita income.

The empirical results of the relationship between regional inequality and regional growth show that the estimated relationship between inequality and growth changes when using different estimation techniques. However, the estimation using panel data regression, the fixed effects and most cases of random effects methods seem to support the theoretical argument and empirical findings that there is a positive correlation between growth and inequality for developing economies like Vietnam's, while pooled OLS regression yields the results of a negative relationship between the two variables: inequality and growth.

Examination of the growth patterns or tests (absolute *β-convergence* test and ADF panel IPS t-bar test based on unit root test) for convergence in regional income was conducted in chapter seven. All the tests for convergence were rejected, which meant that regional income in Vietnam did not converge towards one another. Based on an informal exploratory data analysis we may conclude that divergence of regional income took place in Vietnam, given the size of beta and its high statistical significance.

Parallel with the test for the per capita GDP convergence, the chapter did also attempt to test for the spatial dependence among provinces, in order to explain the growth pattern of one province in relation with that of others. However, the Moran's I and both the LM statistics did not support our prior assumption of the spatial dependence among provinces in the growth progress.

Lastly, chapter eight provided an overview, the conclusions, discussion and recommendations based on the entire research conducted in this study. The chapter drew several key conclusions. First, regarding the regional inequality in health and education, the key point was that the results from the multidimensional approach to measurement and analysis were satisfactory, because generally, welfare inequality in Vietnam, according to the World Bank reports, increased over time, as inequality in income per capita increased. However, the multidimensional approach found that inequality in education and health decreased over the period studied. Additionally, the chapter concluded that spillover effects were important in the case of measuring and analyzing regional inequality for health, and they should be taken into consideration. For education, analyzing the comparative composite indices with respect to education facilities and education enrolments revealed that in some regions these two indices did not coincide, which meant that regions, which were favored in terms of the composite index with respect to education facilities, were not the ones that were favored in terms of the composite index with respect to enrolments and vice versa. This conclusion would be very useful for policymakers in considering the effective way to invest in improving general education in the country.

Chapter eight also described several limitations of the study, regarding the approaches used in the study, the number of indicators used, the sources and duration of the data. However, the limitations were not critical, but if we overcame these setbacks, the results of the study would be more impressive, empirically speaking. In addition, several recommendations for policy were given in chapter eight, regarding policy on health and education, and policy on regional development and sustainable development.

Samenvatting

Deze studie had meerdere doelstellingen. Ten eerste was hij gericht op het onderzoeken van regionale ongelijkheid in Vietnam met behulp van een drietal welvaartsvariabelen: gezondheid en onderwijs, gedurende de periode van 1998-2004 en inkomen per hoofd van de bevolking van 1990 tot 2006. Bovendien is aandacht besteed aan het niveau en de trends van de verschillen tussen- en binnen de regio's voor elk van deze welvaartsvariabelen. De tweede doelstelling van de studie was de relatie te onderzoeken tussen regionale inkomensongelijkheid en regionale economische groei in Vietnam gedurende de periode 1990-2006. De derde en laatste doelstelling van de studie was om de regionale economische groeipatronen in het land te onderzoeken voor de periode 1990-2006. Bij de derde doelstelling werd ook de rol van ruimte en ruimtelijke autocorrelaties bekeken om een beter inzicht te krijgen in de regionale interacties in Vietnam en in het effect op de regionale economische groei. Om deze doelstellingen te bereiken, zijn in de studie verscheidene methoden toegepast om 1) de regionale ongelijkheid van welvaartsvariabelen te meten en te analyseren, 2) de empirische relatie tussen ongelijkheid en groei te onderzoeken en 3) de groeipatronen in de Vietnamese economie te onderzoeken. De studie is onderverdeeld in acht hoofdstukken. Naast de introductie in hoofdstuk een en een afsluitend hoofdstuk acht, dekken hoofdstukken twee tot en met zeven alle bovengenoemde onderwerpen.

Het economisch herstel en de veranderingen in welvaart in Vietnam zijn in detail beschreven en besproken in hoofdstuk twee. De hoofddoelstelling van dit hoofdstuk was om de macro-economische situatie te bespreken, evenals de uitvoering van het economisch herstel, *doi moi* genaamd en het effect daarvan op de welvaart van de bevolking. Zoals besproken in het hoofdstuk, functioneerde het welvaartssysteem in Vietnam voor de *doi moi* heel behoorlijk. De meeste Vietnamezen hadden gelijke toegang tot de meeste collectieve diensten vastgesteld door de overheid, ongeacht geloof, etniciteit of sociale positie. Net zoals in andere communistische landen, was de welvaart voor de *doi moi* redelijk gelijk verdeeld in Vietnam, hoewel de kwaliteit van de geleverde diensten twijfelachtig was. De meeste Vietnamezen hadden vrij toegang tot de gezondheidszorg en scholieren en studenten genoten gratis onderwijs op alle niveaus. Dit resulteerde in een relatief goede volksgezondheid, hoge aantallen studenten en een hoge mate van geletterdheid onder de Vietnamese bevolking vergeleken met andere ontwikkelingslanden met hetzelfde of zelfs iets hoger ontwikkelingsniveau.

Onderwijs en gezondheid na de *doi moi,* zijn ook besproken en vergeleken met de toestand ervoor. Dit is gedaan om het effect te evalueren van relatief hoge economische groeipercentages van 6-9% per jaar op het niveau van de welvaart. Uit deze analyse kwam naar voren dat de regionale ongelijkheid na de *doi moi* was afgenomen.

De onderzoeksmethode met betrekking tot het meten en analyseren van regionale welvaartsongelijkheid is behandeld in hoofdstuk drie. Naast het bespreken van de theoretische benadering en de empirische meetmethoden, voorziet het hoofdstuk ook in een kort overzicht van de literatuur over multidimensionale regionale ongelijkheid. Van de diverse benaderingen voor het meten en analyseren van welvaartsongelijkheid, is in deze studie gekozen voor de multidimensionale benadering, met gebruik van Theil's tweede index als maat voor ongelijkheid. De reden voor het gebruik van Theil's tweede index en de waardevolle eigenschappen van deze index worden ook in dit hoofdstuk besproken. Theil's tweede maat voldoet aan de eigenschap van additieve ontleedbaarheid en is informatief over de regionale welvaartsverdeling. Additieve ontleedbaarheid houdt in dat totale ongelijkheid ontleed kan worden in ongelijkheid binnen en tussen regio's, hetgeen belangrijk is voor het begrijpen van regionale ongelijkheid. Om Theil's tweede ongelijkheidsmaat te berekenen, is het schatten van een samengestelde index nodig. Daarom zijn in dit hoofdstuk *de Principal Component Analysis* (PCA) en *Partial Common Principal Component* (PCPC) technieken uiteengezet.

Het meten en analyseren van regionale ongelijkheid voor gezondheid en onderwijs, zijn beschreven in respectievelijk de hoofdstukken vier en vijf. In hoofdstuk vier zijn de medische voorzieningen en de toestand van de volksgezondheid geanalyseerd. Ook zijn de *spillover*-effecten onderzocht. Zo kunnen medische voorzieningen in een bepaalde provincie een effect hebben op de gezondheid van de mensen in de direct aangrenzende provincie(s). In het algemeen is de regionale ongelijkheid voor beide variabelen enigszins afgenomen in Vietnam in de periode 1998- 2004. De empirische resultaten in hoofdstuk vier laten ook het belang zien van *spillover*- effecten van de gezondheidscomponent. Medische voorzieningen lijken in het geval van contiguïteit meer gelijkelijk verdeeld dan in het geval van non-contiguïteit, aangezien Theil's tweede maat in de eerste situatie lager was dan in de tweede. Dit resultaat geeft aan dat als we het bestaan van *spillover* -effecten zouden negeren, de regionale ongelijkheid overschat zou zijn.

De test voor ruimtelijke afhankelijkheid tussen gezondheidscomponenten is ook uitgevoerd in hoofdstuk vier voor verschillende specificaties van de relatie tussen gezondheid en medische voorzieningen. De resultaten van de testen bevestigden het bestaan van ruimtelijke afhankelijkheid tussen de twee gezondheidscomponenten.

Hoofdstuk vijf behandelt de meting en analyse van de multidimensionale regionale ongelijkheid voor onderwijs. Anders dan in hoofdstuk vier, echter, zijn in dit hoofdstuk de contiguïteitseffecten niet bekeken, omdat de Vietnamese overheid leerlingen niet toestaat onderwijs te volgen in een andere plaats dan waar ze geregistreerd zijn.

Voor regionale ongelijkheid in onderwijsvoorzieningen liet het onderzoek zien dat zowel de 'tussen de regio's' component als de totale ongelijkheid in onderwijsvoorzieningen daalde in de periode 1998-2002 om vervolgens in de periode 2002-2004 weer te stijgen. De ongelijkheid 'binnen de regio' daarentegen steeg tijdens de periode 1998-2002, en daalde in de periode 2002-2004

Voor regionale ongelijkheid in aantallen scholieren stegen zowel de totale regionale ongelijkheid als de ontbonden termen in de eerste periode (1998-2002), terwijl zij in de daaropvolgende periode (2002-2004) daalden. In de drie onderzochte jaren was het aandeel in de totale ongelijkheid van de parameter 'tussen de regio's' groter dan het aandeel van de parameter 'binnen de regio'. Dit houdt in dat de verschillen in aantallen scholieren tussen de regio's belangrijker waren voor de totale ongelijkheid dan de verschillen in aantallen scholieren binnen de regio's.

De comparatieve samengestelde index met betrekking tot onderwijsvoorzieningen en die met betrekking tot aantallen scholieren gepresenteerd in hoofdstuk vijf lieten zien dat er een scherp contrast was tussen de twee onderwijsvariabelen. De regio's met de grootste aantallen scholieren waren regio's in het laagland, terwijl juist daar de minste onderwijsvoorzieningen waren. De regio's met de minste aantallen scholieren lagen in de hoger gelegen gebieden, waar, in termen van de samengestelde index, de meeste onderwijsvoorzieningen waren.

Hoofdstuk zes presenteert de meting en analyse van de regionale inkomens-ongelijkheid en onderzocht de ontwikkeling van inkomensverschillen en de relatie met de economische groei in Vietnamese regio's in de periode van 1990 tot 2006. De resultaten laten zien dat in deze periode de regionale inkomensongelijkheid in het land gematigd is toegenomen. De decompositie van de totale ongelijkheid liet zien dat ongelijkheid 'binnen de regio' afnam, terwijl zij 'tussen de regio's' juist toenam.

Het empirisch onderzoek van de relatie tussen regionale ongelijkheid en regionale groei liet verschillende patronen zien. De schattingen op basis van de paneldataregressie toonden een positieve relatie tussen de mate van economische groei en het beginniveau van inkomen per hoofd van de bevolking in alle onderzochte cases. Gebruikmakend van de *pooled OLS* -techniek, vonden we negatieve relaties tussen groei en ongelijkheid voor alle cases die onderzocht zijn: 1. zonder controles en regionale dummy's, 2. met controles en geen regionale dummy's, en 3. met controles en met regionale dummy's.

Uit de validiteitstest van de methode bleek dat de vaste effectenmethode de voorkeur verdiende boven de aselecte effectenmethode voor de cases: zonder controles en dummy's en met controles maar zonder dummy's in het groeimodel, terwijl de aselecte effecten techniek geschikter was voor de case: met controles en met regionale en tijd dummy's. Het is interessant dat schattingen met beide, de vaste effecten en aselecte effecten, methodes aantoonden dat er zowel een positieve als een negatieve relatie was tussen regionale groei en beginniveau van het inkomen per hoofd.

De empirische bevindingen van de relatie tussen regionale ongelijkheid en regionale groei tonen aan dat de geschatte relatie tussen ongelijkheid en groei afhankelijk is van de gebruikte schattingstechniek. De schatting met behulp van de paneldataregressie, de vaste effecten en de meeste gevallen van de aselecte effectenmethode lijken het theoretische argument en de empirische resultaten te ondersteunen dat er een positieve correlatie is tussen economische groei en ongelijkheid voor ontwikkelingseconomieën zoals Vietnam, terwijl de *pooled OLS* -regressie resulteert in een negatieve relatie tussen de variabelen ongelijkheid en groei.

Onderzoek naar groeipatronen of testen (absolute *β-convergentie* test , de ADF panel IPS t-bar test) voor convergentie in regionaal inkomen werd uitgevoerd in hoofdstuk zeven. Alle testen voor convergentie werden verworpen, wat betekende dat regionaal inkomen in Vietnam niet convergeerde. Op grond van een informele exploratieve data-analyse mogen we concluderen dat divergentie van het regionaal inkomen plaatsvond in Vietnam, gegeven de omvang van de Beta en de hoge statistische significantie ervan.

Parallel aan de test voor convergentie van het regionale gemiddelde inkomen per hoofd (GDP), is in dit hoofdstuk ook getracht te testen op de ruimtelijke afhankelijkheid tussen provincies, om zo het groeipatroon van een provincie te verklaren ten opzichte van de andere. De Moran's I en beide LM-statistieken ondersteunden onze eerdere veronderstelling van ruimtelijke afhankelijkheid tussen provincies in het groeiproces niet.

Ten slotte behandelt hoofdstuk acht het volgende: overzicht, conclusies, discussie en aanbevelingen, gebaseerd op het gehele onderzoek uitgevoerd in deze studie. Het hoofdstuk bevat verscheidene hoofdconclusies. Met betrekking tot de regionale ongelijkheid in gezondheid en onderwijs kon worden aangetoond dat deze is verminderd gedurende de bestudeerde periode. Dit is opmerkelijk omdat in de rapporten van de Wereldbank doorgaans wordt geconcludeerd dat de welvaartsongelijkheid (gemeten als ongelijkheid in het inkomen per hoofd). in Vietnam is toegenomen. Bovendien kan worden geconcludeerd dat *spillover*-effecten belangrijk zijn bij meting van regionale ongelijkheid in gezondheid. Betreffende het onderwijs liet analyse van de comparatieve samengestelde indexcijfers met betrekking tot onderwijsfaciliteiten en inschrijvingen zien dat in sommige regio's deze twee indexcijfers niet overeenstemden, wat betekende dat regio's met een gunstige samengestelde index voor onderwijsvoorzieningen niet die waren die gunstig waren in termen van de samengestelde index met betrekking tot aantallen scholieren en vice versa. Deze conclusie zou nuttig kunnen zijn voor beleidsmakers die op zoek zijn naar effectieve manieren om te investeren in verbetering van het Vietnamese onderwijs.

Appendices to chapter 2

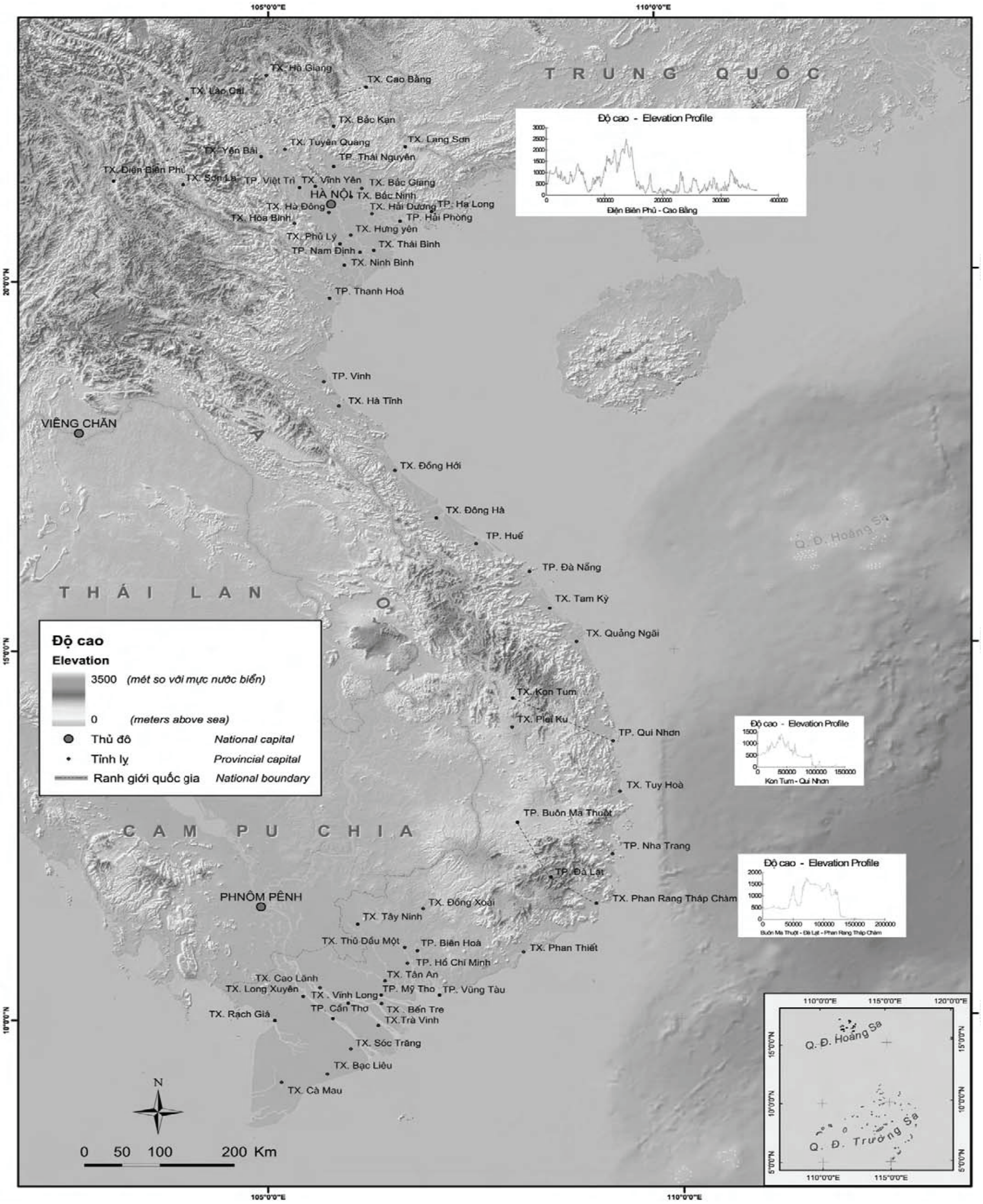
2.1 Names of provinces by region and number of districts in each provinces and regions

Regions/provinces(cities)	Districts	Regions/provinces(cities)	Districts
Red River Delta	**93**	**South Central Coast**	**58**
Hanoi city	12	Da Nang	6
Hai Phong	13	Quang Nam	14
Ha Tay	14	Quang Ngai	13
Hai Duong	12	Binh Dinh	11
Hung yen	10	Phu Yen	7
Ha Nam	6	Khanh Hoa	7
Nam Dinh	10		
Thai Binh	8	**Central Highlands**	**37**
Ninh Binh	8	Kon Tum	7
		Gia Lai	12
North East	**124**	Dak Lak	18
Ha Giang	10		
Cao Bang	11	**South East**	**84**
Lao Cai	11	Ho Chi Minh city	22
Bac Kan	7	Lam Dong	11
Lang Son	11	Ninh Thuan	4
Tuyen Quang	6	Binh Phuoc	6
Yen Bai	9	Tay Ninh	9
Thai Nguyen	9	Binh Duong	7
Phu Tho	12	Dong Nai	9
Vinh Phuc	7	Binh Thuan	9
Bac Giang	10	Ba Ria-Vung Tau	7
Bac Ninh	8		
Quang Ninh	13	**Mekong River Delta**	**107**
		Long An	14
North West	**30**	Dong Thap	11
Lai Chau	10	An Giang	11
Son La	10	Tien Giang	9
Hoa Binh	10	Vinh Long	7
		Ben Tre	8
North Central Coast	**81**	Kien Giang	13
Thanh Hoa	27	Can Tho	8
Nghe An	19	Tra Vinh	8
Ha Tinh	10	Soc Trang	7
Quang Binh	7	Bac Lieu	4
Quang Tri	9	Ca Mau	7
Thua Thien-Hue	9		

2.2 Administrative Map of Vietnam

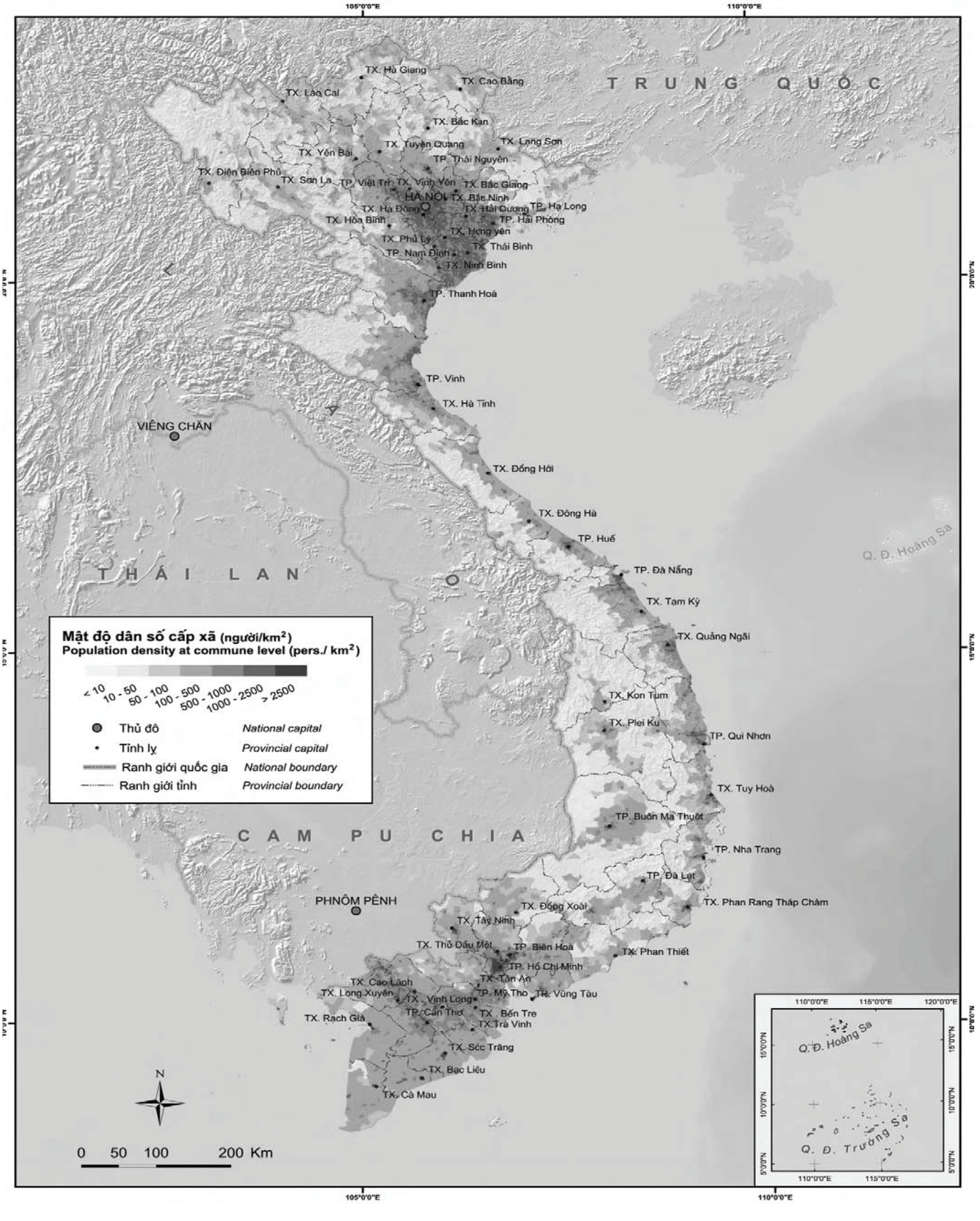
Source: GSO, the Population Census, 1999

Appendix 2.3
Elevation (Relief) Map of Vietnam



Source: GSO, the Population Census, 1999

Appendix 2.4
Population Density Map



Source: GSO, the Population Census, 1999

Appendices to Chapter 3
3.1 The aggregate function of a composite index

In appendix 3.1, we present the aggregate function that forms the composite index of a welfare component under consideration. This representation is drawn from the work by Maasoumi (1986) and Quadrado *et al.* (2001).

In our study, we use a composite index of an underlying unit of welfare under consideration for the calculation of Theil's second measure. Thus a suitable index or aggregator function needs to be defined. As in all index number problems, the intended use of the index affects its ideal formulation. Since we are concerned with the study of the multidimensional distribution of welfare *Xw,* it is reasonable to require the index to have a distribution that is as close as possible to this distribution. The aggregate function is thus defined by Maasoumi as follows.

$$w_{ij} = w\infty\left(\sum_{k=1}^{p} \delta_{jm} s_{ijm}\right)^{-1/\beta} \text{if } \beta \neq 0,-1 \qquad (3.13)$$

where:

δ_{jm} are the weights associated with the coefficients of *m-th* indicators in province *j* of the welfare under consideration. These weights allow the unequal valuation of the different variables so they are comparable to prices for index number. The weights δ_j is defined as follows.

$$\delta_{jm} = s_{jm} / \sum_{m=1}^{p} s_{jm} \qquad (3.14)$$

β is level of aggregation of indicators. β can take the value of -2; -1; -1/2 or 0,

w_{ij} is the composite index of the welfare component under consideration of province *j* in region *i,* and

s_{ijm} is the m- th indicator of welfare under consideration of province *j* of region *i*.

.

Appendix 3.2
The Principal Component Model (PCA)

In this appendix, we present the method of principal component analysis. This is based on the work by Flury (1988) and to some extension on the work by Quadrado (1999). Following Flury, we will present the PCs into two parts such as principal components in the population and principal components of a sample. Finally, we present the interpretation of the eigenvalues and eigenvectors of the principal components.

Principal Components in the Population

We work with a random vector $X = (X_1,.....,X_p)'$. Denote covariance matrix of X by Ψ. For simplicity we assume that $E[X] = 0$. We again assume that the covariance matrix is symmetric and positive definite. The spectral decomposition of Ψ is

$$\Psi = \sum_{j=1}^{p} \lambda_j \beta_j \beta_j' = \beta \Lambda \beta' \qquad (3.15)$$

where $\beta = (\beta_1,....,\beta_p)$ is an orthogonal $p \times p$ matrix of the eigenvectors of the covariance matrix Ψ, which are associated with the corresponding eigenvalues or characteristic roots λ_j (*j = 1,.....,p*) of the diagonal matrix Λ. We assume that the characteristic roots λ_j are arranged in descending order that means $\lambda_1 \geq \lambda_2 \geq ..... \geq \lambda_p$. From the orthogonality of β it follows that $\beta' \Psi \beta = \Lambda$. According to Flury the variance of any normalized linear combination of *X* cannot exceed largest eigenvalue λ_1

$$\text{var}(\beta_1' X) = \beta_1' \Psi \beta_1 = \lambda_1 \qquad (3.16)$$

and the first principal component of X is defined as

$$U_1 = \beta_1' X \qquad (3.17)$$

And the covariance of X is equal to Λ. Therefore, the principal components of *X* (denotes as *U* with $U_h = \beta_h' X = (U1,.....,Up)$) are unrelated to λ_h *(h =1,.....,p).*

It follows that the first principal component in (3.17) is a normalized linear combination with maximum variance equals λ_1. In this fashion, the second component is also a normalized linear combination with largest variance of λ_2 but uncorrelated with the first principal component. This fashion will be followed to *p* principal components with component *j-th* which has a variance of λ_j.

Summarising this procedure, the principal components of *X* are defined as the *p*-variate random vector

$$U = \begin{bmatrix} U_1 \\ \vdots \\ U_p \end{bmatrix} = \beta' X \qquad (3.18)$$

where the columns of β are the normalized characteristic vectors of Ψ. The covariance matrix of *U* is

$$Cov(U) = E[UU'] = \beta' \Psi \beta = \Lambda \qquad (3.19)$$

which implies that the principal components are pairwise uncorrelated. Furthermore, principal component U_h has maximum variance among all normalized linear combinations that are uncorrelated with U_1 to U_{h-1}.

The principal components have quite appealing properties. For instance, the total variance of X (δ^2_{total}) and the generalized variance of X (δ^2_{gen}) are defined as

$$\delta^2_{total} = \sum_{i=1}^{p} \text{var}[X] = tr\Psi \quad (3.20)$$

and $$\delta^2_{gen} = \det \Psi \quad (3.21)$$

respectively, where "*tr*" and "*det*" denote the trace and the determinant of the random vector *X*. These two measures of multivariate dispersion are invariant under the principal component transformation,

$$\delta^2_{total} = tr(\beta\Lambda\beta') = tr(\Lambda\beta'\beta) = tr\Lambda = \sum_{j=1}^{p} \lambda_j = \sum_{j=1}^{p} \text{var}[U_j] \quad (3.22)$$

and $$\delta^2_{gen} = \det(\beta\Lambda\beta') = \det(\beta)\det(\Lambda)\det(\beta') = \det(\Lambda)\det(\beta'\beta) = \det\Lambda \quad (3.23)$$

Principal components enjoy a number of intuitively appealing properties (Flury, 1988; MacCabe,1984). Flury had mentioned some of those properties involving in a maximization or minimization over $q(<p)$ linear functions of the p-variate random vector *X*, subject to orthogonality constraints. The formal transformation takes the form:

$$Y = A'X \quad (3.24)$$

where a is a p x q matrix such that $A'A = I_q$. If we take the characteristic vectors associated with the q largest characteristic roots of $\Psi = Cov(X)$ to form columns of Λ. The covariance matrix of Y can be written as

$$\Psi_Y = \Lambda'\Psi\Lambda \quad (3.25)$$

with characteristic roots of Ψ stacking in the descending order as usual as $\lambda_1 \geq \lambda_2 \geq \ldots. \geq \lambda_p$.

We now work with q principal components where their total and generalized variance presented in equations (3.20) and (3.21) are to be maximized as follows:

$$\max\ tr(\Psi_Y) = \sum_{j=1}^{q} \lambda_j \quad (3.26)$$

and $$\max\ \det(\Psi_Y) = \prod_{j=1}^{q} \lambda_j \quad (3.27)$$

respectively. To continue, we assume that X_1 and X_2 are independent p-variate random vectors with the same distribution as *X*, and $Y_h = A'X'_h$ $(h = 1,2)$. Total and generalized variances can be re-written as follows:

$$\max\ E\left[(Y_1 - Y_2)'(Y_1 - Y_2)\right] = 2\sum_{j=1}^{q} \lambda_j \quad (3.28)$$

and $$\max \det \left\{ E\left[\left(Y_1 - Y_2\right)\left(Y_1 - Y_2\right)' \right] \right\} = 2\prod_{j=1}^{q} \lambda_j \tag{3.29}$$

The last two expressions in equations (3.28) and (3.29) of maximization of total and generalized variances are based on the assumption that X has a multivariate normal distribution with mean vector 0.

The principal Components of a Sample

Statistically, there is little difference between the principal components of a population and the principal components of a sample. Let us denote S as a covariance matrix of a random sample $X=X_1,\ldots,X_n$ of size N = n + 1 with multivariate normal distribution $N(\mu, \Psi)$ in its unbiased estimator $\hat{\Psi}$, where S can be expressed as:

$$S = \frac{1}{N-1}\sum_{h=1}^{N}(X_h - \overline{X})(X_h - \overline{X})' \tag{3.30}$$

and mean $\overline{X}$ as $\overline{X} = \frac{1}{N}\sum_{h=1}^{N} X_h$ and assume that covariance matrix $\Psi = \beta\Lambda\beta'$ is positive definite. Then $S^* = (N-1)S/N$ is the maximum likelihood estimate of Ψ. Furthermore, if $N > p$, then with probability of 1, S is positive definite, and the p characteristic roots of S are distinct. The spectral decomposition of S is

$$S = BLB' = \sum_{j=1}^{p} l_j b_j b_j' \tag{3.31}$$

where $B = (b_1,\ldots,b_p)$ is orthogonal eigenvectors, and $L = diag(l_1,\ldots,l_p)$. Assuming that all characteristic roots of Ψ are distinct, the spectral decomposition of Ψ is unique. According to the theorem on the invariance of maximum likelihood estimates that the maximum likelihood estimates of β and Λ are the characteristic vectors and roots of S*, and that this can be expressed as

$$\hat{\beta} = B$$
$$\hat{\Lambda} = \frac{N-1}{N}L \tag{3.32}$$

According to the theory of maximum likelihood estimation, it is obvious that B and L are the consistent estimators of β and Λ.

Due to usefulness of the computation of some measure of variability of the parameter estimates whenever the parameters of a statistical model are estimated, we now turn to discuss about the inference based on the principal components characteristic roots and coefficients.

Inference for characteristic roots

For sufficiently large sample size n we can treat *j-th* root l_j as approximately

$$l_j \sim N(\lambda_j, 2\lambda_j^2 / n) \tag{3.33}$$

and the consistent estimate of the standard error of l_j is defined as

$$s(l_j) = \sqrt{\frac{2}{n} l_j} \quad (3.34)$$

A large sample confident interval covering λ_j with probability approximately $1\text{-}\alpha$ is given by

$$\left[l_j \left(1 - z_{\alpha/2} \sqrt{2/n}\right), l_j \left(1 + z_{\alpha/2} \sqrt{2/n}\right) \right] \quad (3.35)$$

In principal components, the relative contribution of one or several components to the trace of the covariance matrix is very important. When we discard the last *p-q* components the relative contribution to the trace of the covariance matrix of these discarded components is computed by

$$f = \frac{\lambda_{q+1} + \ldots.. + \lambda_p}{\lambda_1 + \ldots\ldots + \lambda_p} \quad (3.36)$$

if the value of f in equation (3.36) is sufficiently small, say smaller than a fixed value f_o, then the last *p-q* components can be discarded.

To test the hypothesis $f \le f_o$ for a fixed $f_o \in (0,1)$, versus the alternative $f > f_o$, we can use z(f_o) and reject the null if z(f_o) is larger than z_α, the upper α quantile of the standard normal distribution. The test is based on

$$\hat{f} = \frac{\sum_{j=q+1}^{p} j_j}{trS} \quad (3.37)$$

and the hypothesis is accepted if

$$\hat{f} \le f_0 + \sqrt{\frac{2}{n}} z_\alpha \left[f_0^2 \sum_{j=1}^{q} l_j^2 + (1 - f_0)^2 \sum_{j=q+1}^{p} l_j^2 \right]^{1/2} \quad (3.38)$$

Inference for principal component coefficients (eigenvectors)

When principal components are interpreted, it is important to be aware that the estimated coefficients b_{jh} have some sampling variability. Therefore, it is necessary to compute standard errors of estimated principal component coefficients. These errors are based on the assumption that the respective principal components are well defined in the population. Before interpreting a principal component, we should be certain that the associated characteristic roots λ_h of the covariance matrix Ψ are distinct. Therefore, it is necessary to check the assumption $\lambda_{h-1} > \lambda_h > \lambda_{h+1}$, where λ_h is the eigenvalues associated with eigenvectors of the coefficients $b_h = (b_{1h}, \ldots.., b_{ph})'$. Checking the assumption can only be done by using the sphericity test for the null hypothesis $\lambda_{h-1} = \lambda_h$ or $\lambda_h = \lambda_{h+1}$ versus the alternative that these eigenvalues are distinct. The sphericity statistic for the former test is

$$S(l_{h-1}, l_h) = 2n \log \frac{l_{h-1} + l_h}{\sqrt{l_{h-1} l_h}} \quad (3.39)$$

which follows asymptotic distribution as chi square with 2 degrees of freedom under the null hypothesis. When the null is rejected, asymptotic standard errors of coefficients can be computed using

$$S(b_{mh}) = \left[\frac{1}{n} l_h \sum_{\substack{j=1 \\ j \neq h}}^{p} \frac{l_j}{(l_j - l_h)^2} b_{mj}^2 \right]^{1/2} \tag{3.40}$$

In case the null is not rejected, the component coefficients are mathematically not uniquely defined. The associated eigenvalues only differ by sampling error. Therefore, there is no need to compute standard errors of the coefficients to interpret these coefficients.

Appendix 3.3
The Partial Common Principal Component (PCPC) Model

A brief outline of the PCPC model presented here is drawn from the work by Flury (1988), Phillips (1997) with some extension and explanation for this case of study.

Consider *t (t = 1,….,k)* samples, i.e. in the present case, the samples are years of surveys. The basic assumption of PCPC is that a subset of principal components is common for all samples whereas the remaining principal components are specific to each sample. In the present case of this study *k=3*, the 1998 VLSS survey, 2002 VHLSS survey and 2002 VHLSS survey. Covariance matrices of all samples could share one, two or all components, and they could be equal, proportional or unrelated to each other. According to Phillips (1997), the essence of the PCPC model is to compare the structure of two or more covariance matrices in a hierarchy fashion. A complete hierarchy of relationships among covariance matrices is as follows.

1. Covariance matrices of all samples are identical.
2. These covariance matrices have the same eigenvectors but their eigenvalues are different by a proportional constant. Therefore each element of the covariance matrix is multiplied by a single constant.
3. All these covariance matrices share common principal components but their eigenvalues differ and are specific in each sample (each covariance matrix).
4. Covariance matrices share *q* components ($q<p$) (*p* is number of variables or indicators of the welfare component under consideration) and the remaining eigenvectors (*p-q*) are specific in each sample.
5. There is no relationship among covariance matrices.

According to Flury, the PCPC model can be presented as follows.

Formally, the hypothesis of the partial common principal component (CPC) (of order *q*) is

$$H_{pcpc}(q) : \Psi_t = \beta^{(t)} \Lambda_t \beta^{(t)'}, \quad t=1,\ldots,k, \tag{3.41}$$

where:

q: number of components is considered, $q < p$ where *p* is number of variables or indicators,

Ψ_t are positive definite symmetric *(p x p)* matrices,

$$\Lambda_t = diag(\lambda_{t1},....,\lambda_{tp}),\tag{3.42}$$

where λ_t are eigenvalues associated with t^{th} sample,

and $$\beta^{(t)} = (\beta_c, \beta_s^{(t)})\tag{3.43}$$

where all $\beta^{(t)}$ are orthogonal (*p x p)* matrices. The part β_c of dimension (*p x q*) are the common eigenvectors for all samples, and the parts $\beta_s^{(t)}$ of dimension (*p x (p-q))* are specific eigenvectors for each sample. So we can present β_c and $\beta_s^{(t)}$ as:

$$\beta_c = (\beta_1,......,\beta_q)\tag{3.44}$$

and $$\beta_s^{(t)} = (\beta_{(q+1)}^{(t)},......,\beta_p^{(t)}), \qquad t = 1,....,k\tag{3.45}$$

As these matrices are orthogonal, $H_{pcpc}(p-1)$ means $H_{pcpc}(p)$, which is the ordinary PCP model. So we restrict the number of common components (*q*) in the partial model to the range $1 \le q \le p-2$. Therefore, the partial CPC model requires the number of variable *p* of at least 3.

Using the notation in equations (3.19) and (3.20), the partial CPC model can be written in spectral decomposition form as follows.

$$\Psi_i = \sum_{j=1}^{q} \lambda_{tj}\beta_j\beta_j' + \sum_{j=q+1}^{p} \lambda_{tj}\beta_j^{(t)}\beta_j^{(t)'}, \qquad t = 1,......,k\tag{3.46}$$

where the vectors $\beta_1,....,\beta_q$ demote the common characteristic vectors (eigenvectors) of all covariance matrices (*k*), and $\beta_{q+1}^{(t)},......,\beta_p^{(t)}$ are the specific characteristics vectors of these covariance matrices.

Applying this model to the present study, we consider to fit the partial CPC model with one common component (*q=1*) which is associated with the highest eigenvalue. Accordingly, this component recovers the most variation in all samples. Therefore, the composite index calculated can explain the maximum information contained in the raw data. Therefore, the hypothesis of one common component is restricted to the eigenvector that retains the highest proportion of variance. According to Flury (1997), we assume cardinal rank order of the eigenvalues of all samples is as

$$\lambda_{t1\ge},............,\ge \lambda_{tp}\tag{3.47}$$

were λ_t are eigenvalues associated with t^{th} sample.

To estimate the partial CPC model we apply the Maximum Likelihood Estimation method. This method is based on the *k* independent covariance matrices S_t with Wishart distribution of n_t *d*egrees of freedom and parameter matrix Ψ_t / n_t. The joint likelihood function of $\Psi_1,......,\Psi_k$, given covariance matrices $S_1,......,S_k$ is as below.

$$L(\Psi_1,......,\Psi_k) = C \times \prod_{t=1}^{k} etr(-\frac{n_t}{2}\Psi_t^{-1}S_t)(\det \Psi_t)^{-n_i/2}\tag{3.48}$$

where *C* is a constant and does not depend on the parameters, n_t is the sample size for *t-th* sample (in our case is the survey), *etr* stands for the exponential function of the trace of covariance matrices, and *det* is the determinants of the matrix Ψ_t. Maximizing the likelihood of the equation (3.19) is equivalent to minimizing the function:

$$g(\Psi_1,......,\Psi_k) = -2\log L + 2\log C = \sum_{t=1}^{k} n_t \left[\log(\det \Psi_t) + tr(\Psi_t^{-1} S_t)\right] \qquad (3.49)$$

Writing $\beta = (\beta_1,....,\beta_p)$ and assuming that hypothesis in equation (3.16) on common component $H_{pcpc}(q)$ holds for a fixed $q(1 \le q \le p-2)$, equation (3.24) can be transformed into equation

$$g(\beta_c, \beta_s^1,......,\beta_s^k, \Lambda_1......,\Lambda_k) = \sum_{t=1}^{k} n_t \left[\sum_{j=1}^{p} \log \lambda_{tj} + \sum_{j=1}^{q} \frac{\beta_j' S_t \beta_t}{\lambda_{tj}} + \sum_{j=q+1}^{p} \frac{\beta_j^{(t)'} S_t \beta_j^{(t)}}{\lambda_{tj}}\right] \qquad (3.50)$$

This function is to be minimized under the constraint of orthogonality of all eigenvectors *($\beta^{(i)}$)* of the *k* covariance matrices. According to Flury (1998), the orthogonality is formalized as follows

$$\beta_h' \beta_j = \begin{cases} 0, & if\ h \ne j \\ 1, & if\ h = j \end{cases} 1 \le h, j \le q \qquad (3.51)$$

$$\beta_h^{(t)'} \beta_j^{(t)} = \begin{cases} 0, & if\ h \ne j \\ 1, & if\ h = j \end{cases} q \le h, j \le p, t = 1,...,k \qquad (3.52)$$

$$\beta_h' \beta_j^{(t)} = 0, \qquad t = 1,.......,k, \quad 1 \le h \le q \le j \le p \qquad (3.53)$$

By carrying out manipulations of the equation (3.46) to equation (3.53) the maximum likelihood estimates for eigenvectors and coefficients can be obtained from equations (Flury,1988)

$$\lambda_{tj} = \begin{cases} \beta_j' S_t \beta_j, & j = 1,.....,q \\ \beta_j^{(t)'} S_t \beta_j^{(t)}, & j = q+1,...., p \end{cases} t = 1,....,k \qquad (3.54)$$

$$\beta_h^{(t)'} S_t \beta_j^t = 0, \quad h \ne j \qquad (3.55)$$

$$\beta_m' \left(\sum_{t=1}^{k} \frac{\lambda_{tm} - \lambda_{th}}{\lambda_{tm}\lambda_{th}} n_t S_t \right) \beta_h = 0, \quad 1 \le m, h \le q, m \ne h \qquad (3.56)$$

Both, equation (3.54) and (3.56) are valid for m, h ≤ q for the common components.
And also from Flury (1988) we get

$$\left(\frac{1}{\lambda_{rj}} - \frac{1}{\lambda_{rm}} \right) n_r \beta_m' S_r \beta_j^{(r)} = \beta_j^{(r)'} \left[\sum_{t=1; t \ne r}^{k} \left(\frac{n_t S_t \beta_m}{\lambda_{tm}} - \sum_{h=q+1}^{p} \delta_{mh}^t \beta_h^{(t)} \right) \right] \qquad (3.57)$$

r=1,….,k; 1 ≤ m ≤ q ≤ j ≤ p

where $$\delta_{(mh)}^t = n_t \beta_t' S_t \beta_h^{(t)} / \lambda_{th} \qquad (3.58)$$

are the Lagrange multipliers for the restrictions. The equation (3.57) links the common and the specific components. Thus the likelihood equations consist of (3.54), (5.55), (3.56), (3.57), (3.58) to be solved under the orthogonality constraints (3.51) to (3.53).

Following Flury (1988), the test statistic for an exact and an approximate maximum log-likelihood test of the null hypothesis of a common eigenvalue versus the alternative of unrelated covariance matrices (as we denote exact maximum likelihood estimates by $\hat{\beta}^{(t)}$, $\hat{\Lambda}_t$, and $\hat{\Psi}_t$) by

$$\chi^2_{pcpc}(q) = -2\log\frac{L\left(\hat{\Psi}_1,\ldots..,\hat{\Psi}_k\right)}{L(S_1,\ldots...,S_k)} = \sum_{t=1}^{k} n_t \log\frac{\det\hat{\Psi}_t}{\det S_t} \tag{3.59}$$

and

$$\overline{\chi}^2_{pcpc}(q) = -2\log\frac{L(\overline{\Psi}_1,\ldots...,\overline{\Psi}_k)}{L(S_1,\ldots..,S_k)} = \sum_{t=1}^{k} n_t \log\frac{\det\overline{\Psi}_t}{\det S_t} \tag{3.60}$$

respectively**.** The number of parameters estimated under null hypothesis is *p(p-1)/2 + kp + (k-1)(p-1)(p-q-1)/2*. Under the alternative, the number of parameters is *kp(p+1)/2*. Thus, the asymptotic null distribution of χ^2 is chi square with *(k-1)q(2p-q-1)/2* degrees of freedom.

If the test for the null hypothesis of a common eigenvalue is not rejected, the weights attached to the indicators will be the same for all the samples, i.e. the common weights are arrived. They are the coefficients of the first component estimated. If the hypothesis is rejected, it means that the PCPC model does not fit the data, so the PCPC model is applied to each sample separately. Then the weights associated to variables are different for each sample and coincide to the first principal components of each sample separately.

The stability of the coefficients of the first common component can be investigated by considering the value of the standard errors. According to Flury (1988), if a standard error of a coefficient is smaller than 0.1, then that coefficient may be presumed to be stable.

In this case, the standard errors of the first eigenvectors of the common principal component under partial CPC are computed using consistent estimates as

$$\hat{\theta}^{(t)}_{jh} = r_t^{-1}\frac{\hat{\lambda}_{tj}\hat{\lambda}_{th}}{(\lambda_{tj}-\lambda_{th})^2} \tag{3.61}$$

where $r_t = {}^{n_t}\!/_{n}$ where *n* is numbers *k* of samples (groups), $n=n_1+\ldots..+n_k$.

and
$$\hat{\theta}_{jh} = \left(\sum_{i=1}^{k}\hat{\theta}^{(t)^{-1}}_{jh}\right)^{-1} \tag{3.62}$$

standard error of $\hat{\beta}_{mh}$ is calculating using

$$s\left(\hat{\beta}_{mh}\right) = \left(\frac{1}{n}\sum_{\substack{j=1 \\ j \neq h}}^{p} \hat{\theta}_{jh}\hat{\beta}_{mj}^{2}\right)^{1/2} \qquad (3.63).$$

Appendices to Chapter 4
4.1 The administrative Division in Vietnam

No.	Code	Administrative Division	No.	Code	Administrative Division
I	**1**	**Red River Delta**	**V**	**5**	**South Central Coast**
1	101	Hanoi	32	501	Da Nang
2	102	Hai Phong	33	502	Quang Nam
3	103	Vinh Phuc	34	503	Quang Ngai
4	104	Ha Tay	35	504	Binh Dinh
5	105	Bac Ninh	36	505	Phu Yen
6	106	Hai Duong	37	506	Khanh Hoa
7	107	Hung Yen	**VI**	**6**	**Central Highlands**
8	108	Ha Nam			
9	109	Nam Dinh	38	601	Kon Tum
10	110	Thai Binh	39	602	Gia Lai
11	111	Ninh Binh	40	603	Dak Lak
II	**2**	**North East**	41	604	Lam Dong
			VII	**7**	**South East**
12	201	Ha Giang			
13	202	Cao Bang	42	701	Ho Chi Minh City
14	203	Lao Cai	43	702	Ninh Thuan
15	204	Bac Kan	44	703	Binh Phuoc
16	205	Lang Son	45	704	Tay Ninh
17	206	Tuyen Quang	46	705	Binh Duong
18	207	Yen Bai	47	706	Dong Nai
19	208	Thai Nguyen	48	707	Binh Thuan
20	209	Phu Tho	49	708	Ba Ria – Vung Tau
21	210	Bac Giang	**VIII**	**8**	**Mekong River Delta**
22	211	Quang Ninh			
III	**3**	**North West**	50	801	Long An
			51	802	Dong Thap
23	301	Lai Chau	52	803	An Giang
24	302	Son La	53	804	Tien Giang
25	303	Hoa Binh	54	805	Vinh Long
IV	**4**	**North Central Coast**	55	806	Ben Tre
			56	807	Kien Giang
26	401	Thanh Hoa	57	808	Can Tho
27	402	Nghe An	58	809	Tra Vinh
28	403	Ha Tinh	59	810	Soc Trang
29	404	Quang Binh	60	811	Bac Lieu
30	405	Quang Tri	61	812	Ca Mau
31	406	Thua Thien - Hue			

Appendix 4.2
Travel inputs in 1st order contiguous provinces measured in terms of distance between provincial capitals

Province	Neighboring Provinces	Distance (km)	Province	Neighboring Provinces	Distance (km)
1. Lai Chau	Lao Cai	175		Quang Tri	90
	Son La	100	30. Quang Tri	Quang Binh	90
2. Lao Cai	Ha Giang	120		Thua Thien-H	64
	Yen Bai	140	31. ThuaThien-H	Quang Tri	64
	Son La	135		Da Nang	85
	Lai Chau	175		Quang Nam	146
3. Son La	Lao Cai	135	32. Da Nang	Thua Thien-H	85
	Yen Bai	130		Quang Nam	62
	Phu Tho	160	33. Quang Nam	Thua Thien-H	146
	Hoa Binh	170		Da Nang	62
	Thanh Hoa	275		Kon Tum	150
4. Yen Bai	Ha Giang	125		Quang Ngai	63
	Lao Cai	140	34. Kon Tum	Quang Nam	150
	Tuyen Quang	35		Quang Ngai	125
	Son La	130		Gia Lai	46
	Phu Tho	75	35. Quang Ngai	Quang Nam	63
5. Ha Giang	Lao Cai	120		Kon Tum	125
	Yen Bai	125		Gia Lai	160
	Tuyen Quang	119		Binh Dinh	163
	Cao Bang	135	36. Gia Lai	Kon Tum	46
6.Tuyen Quang	Ha Giang	119		Quang Ngai	160
	Yen Bai	35		Binh Dinh	142
	Phu Tho	62		Phu Yen	131
	Vinh Phuc	70		Dak Lak	150
	Thai Nguyen	75	37. Binh Dinh	Gia Lai	142
	Bac Kan	78		Quang Ngai	163
	Cao Bang	150		Phu Yen	64
7. Cao Bang	Ha Giang	135	38. Phu Yen	Dak Lak	150
	Tuyen Quang	150		Gia Lai	131
	Bac Kan	75		Binh Dinh	64
	Lang Son	105		Khanh Hoa	96
8. Phu Tho	Son La	160	39. Dak Lak	Gia Lai	150
	Yen Bai	75		Phu Yen	150
	Tuyen Quang	62		Khanh Hoa	140
	Vinh Phuc	21		Lam Dong	96
	Ha Tay	52		Binh Phuoc	188
	Hoa Binh	53	40. Khanh Hoa	Lam Dong	94
9. Bac Kan	Tuyen Quang	78		Ninh Thuan	77
	Cao Bang	75		Dak Lak	140
	Lang Son	101		Phu Yen	96
	Thai Nguyen	60	41. Lam Dong	Dak Lak	96
10.Thai Nguyen	Tuyen Quang	75		Khanh Hoa	94
	Bac Kan	60		Ninh Thuan	75

	Lang Son	103		Binh Thuan	122
	Bac Giang	52		Dong Nai	217
	Hanoi	65		Binh Phuoc	183
	Vinh Phuc	41	42. Ninh Thuan	Binh Thuan	126
11. Vinh Phuc	Phu Tho	21		Lam Dong	75
	Tuyen Quang	70		Khanh Hoa	77
	Thai Nguyen	41	43. Binh Thuan	Ba Ria-V.Tau	140
	Hanoi	31		Dong Nai	146
	Ha Tay	43		Lam Dong	122
12. Lang Son	Cao Bang	105		Ninh Thuan	126
	Bac Kan	101	44. Binh Phuoc	Binh Duong	66
	Thai Nguyen	103		Tay Ninh	88
	Bac Giang	87		Dak Lak	188
	Quang Ninh	107		Lam Dong	183
13. Bac Giang	Thai Nguyen	52		Dong Nai	65
	Lang Son	87	45. Dong Nai	Binh Phuoc	65
	Quang Ninh	102		Lam Dong	217
	Hai Duong	40		Binh Thuan	146
	Bac Ninh	19		Ba Ria-V.Tau	75
	Hanoi	60		Ho Chi Minh	25
14. Hung Yen	Thai Binh	36		Binh Duong	18
	Ha Nam	19	46. Ba Ria-V.Tau	Ho Chi Minh	68
	Ha Tay	45		Dong Nai	75
	Hanoi	51		Binh Thuan	140
	Bac Ninh	57	47. Ho Chi Minh	Long An	46
	Hai Duong	42		Tien Giang	63
15. Ha Tay	Hoa Binh	53		Tay Ninh	92
	Phu Tho	52		Binh Duong	27
	Vinh Phuc	43		Dong Nai	25
	Hanoi	11		Ba Ria-V.Tau	68
	Hung Yen	45	48. Tay Ninh	Binh Phuoc	88
	Ha Nam	50		Binh Duong	74
16. Hoa Binh	Thanh Hoa	129		Ho Chi Minh	88
	Son La	170		Long An	95
	Phu Tho	53	49. Binh Duong	Ho Chi Minh	27
	Ha Tay	53		Tay Ninh	74
	Ha Nam	81		Binh Phuoc	66
	Ninh Binh	96		Dong Nai	18
17. Bac Ninh	Hanoi	29	50. Long An	Tay Ninh	95
	Bac Giang	19		Ho Chi Minh	46
	Hai Duong	37		Tien Giang	24
	Hung Yen	70		Dong Thap	78
18. Quang Ninh	Hai Phong	46	51. Tien Giang	Ben Tre	15
	Hai Duong	77		Vinh Long	46
	Bac Giang	102		Dong Thap	82
	Lang Son	107		Long An	24
19. Hai Phong	Quang Ninh	46	52. Dong Thap	An Giang	29
	Hai Duong	42		Long An	78

	Thai Binh	58		Tien Giang	82
20. Hai Duong	Hung yen	42		Vinh Long	45
	Bac Ninh	37		Can Tho	49
	Bac Giang	40	53. Ben Tre	Tra Vinh	38
	Quang Ninh	77		Vinh Long	47
	Hai Phong	42		Tien Giang	15
	Thai Binh	56	54. Vinh Long	Tra Vinh	55
21. Hanoi	Ha Tay	11		Soc Trang	70
	Vinh Phuc	31		Can Tho	57
	Thai Nguyen	65		Dong Thap	45
	Bac Giang	60		Tien Giang	46
	Bac Ninh	29		Ben Tre	47
	Hung Yen	51	55. Tra Vinh	Soc Trang	55
22. Thai Binh	Hai Phong	58		Can Tho	65
	Hai Duong	56		Vinh Long	55
	Hung Yen	36		Ben Tre	38
	Ha Nam	50	56. An Giang	Dong Thap	29
	Nam Dinh	21		Can Tho	55
23. Ha Nam	Ha Tay	50		Kien Giang	54
	Hung Yen	19	57. Can Tho	Kien Giang	79
	Thai Binh	50		An Giang	55
	Nam Dinh	28		Dong Thap	49
	Ninh Binh	32		Vinh Long	57
	Hoa Binh	81		Soc Trang	53
24. Nam Dinh	Ninh Binh	30		Bac Lieu	86
	Ha Nam	28	58. Soc Trang	Bac Lieu	51
	Thai Binh	21		Can Tho	53
25. Ninh Binh	Thanh Hoa	56		Vinh Long	70
	Hoa Binh	96		Tra Vinh	55
	Ha Nam	32	59. Kien Giang	An Giang	54
	Nam Dinh	30		Can Tho	79
26. Thanh Hoa	Hoa Binh	129		Bac Lieu	110
	Ninh Binh	56		Ca Mau	96
	Nghe An	129	60. Bac Lieu	Ca Mau	68
27. Nghe An	Thanh Hoa	129		Kien Giang	110
	Ha Tinh	43		Can Tho	86
28. Ha Tinh	Nghe An	43		Soc Trang	51
	Quang Binh	127	61. Ca Mau	Kien Giang	96
29. Quang Binh	Ha Tinh	127		Bac Lieu	110

Appendix 4.3
Transformed health data: the contiguity and non-contiguity case
4.3.1 Transformed health data for the 1998 sample

No	Province	Health Facilities								Health Status	
		Childhood immunization		Hospital Beds		Medical Doctors		Pharmacist		Life Expectancy	Infant Mortality
		*NonCon.	** Con.	*Non Con.	** Con.	*Non Con.	** Con.	*Non Con.	** Con.		
1	Hanoi	1	1	.388	.416	.55	.605	1	1	.757	.991
2	Hai Phong	.999	.927	.270	.284	.60	.628	.795	.771	.734	.985
3	Vinh Phuc	.998	.969	.161	.193	.30	.352	.623	.643	.722	.982
4	Ha Tay	.994	.994	.194	.228	.30	.355	.436	.489	.699	.973
5	Bac Ninh	.981	.947	.168	.193	.46	.499	.605	.619	.714	.976
6	Hai Duong	.999	.979	.227	.253	.35	.405	.588	.609	.737	.982
7	Hung Yen	.998	1	.174	.210	.33	.391	.290	.365	.720	.981
8	Ha Nam	.993	.988	.278	.307	.35	.406	1	.997	.740	.983
9	Nam Dinh	.979	.957	.184	.213	.47	.519	.688	.720	.753	.978
10	Thai Binh	1	.976	.266	.292	.52	.568	.644	.671	.765	.977
11	Ninh Binh	.989	.927	.214	.233	.42	.448	.599	.602	.714	.974
12	Ha Giang	.900	.815	.254	.263	.25	.265	.448	.432	.588	.934
13	Cao Bang	.916	.832	.302	.311	.58	.595	.611	.586	.624	.947
14	Lao Cai	.985	.886	.284	.291	.42	.430	.404	.389	.660	.961
15	Bac Kan	.926	.854	.339	.352	.53	.557	.509	.502	.666	.963
16	Lang Son	.971	.892	.205	.218	.52	.545	.766	.739	.652	.941
	Tuyen		.907		.306		.437		.401		
17	Quang	.957		.284		.40		.385		.668	.968
18	Yen Bai	.876	.809	.288	.301	.51	.527	.633	.609	.672	.966
	Thai		.900		.305		.599		.676		
19	Nguyen	.939		.281		.56		.669		.696	.971
20	Phu Tho	.976	.924	.193	.214	.36	.391	.380	.395	.710	.979
21	Bac Giang	.980	.946	.178	.203	.33	.380	.523	.551	.674	.966
22	Quang Ninh	.946	.871	.217	.228	.51	.532	.468	.465	.711	.977
23	Lai Chau	.980	.870	.197	.201	.34	.345	.442	.417	.636	.949
24	Son La	.887	.802	.297	.305	.27	.282	.613	.583	.661	.961
25	Hoa Binh	.907	.846	.239	.254	.40	.420	.449	.451	.673	.965
26	Thanh Hoa	.968	.874	.325	.332	.33	.341	.539	.516	.694	.968
27	Nghe An	.941	.845	.216	.223	.26	.267	.497	.473	.706	.972
28	Ha Tinh	.997	.894	.239	.243	.29	.296	.418	.398	.721	.972
29	Quang Binh	.970	.865	.155	.160	.20	.207	.340	.322	.679	.967
30	Quang Tri	.885	.798	.245	.249	.45	.458	.297	.282	.685	.954
	Th. Thien -		.890		.192		.426		.161		
31	Hue	.988		.184		.41		.163		.716	.976
32	Da Nang	.975	.878	.366	.372	.80	.811	.526	.495	.756	.986
33	Quang Nam	.916	.836	.209	.220	.40	.419	.226	.221	.698	.975

34	Quang Ngai	.883	.800	.191	.199	.29	.301	.345	.328	.698	.967
35	Binh Dinh	.989	.886	.164	.168	.32	.327	.411	.389	.700	.970
36	Phu Yen	.966	.877	.161	.168	.39	.402	.368	.353	.661	.962
37	Khanh Hoa	.976	.889	.149	.156	.37	.385	.495	.470	.723	.971
38	KonTum	.943	.848	.390	.396	.43	.438	.159	.155	.572	.934
39	Gia Lai	.910	.830	.244	.253	.25	.264	.226	.221	.618	.953
40	Dak Lak	.907	.824	.186	.193	.37	.382	.113	.115	.656	.952
41	Lam Dong	.930	.855	.213	.221	.41	.426	.361	.348	.707	.977
42	Ho Chi Minh	.802	.795	.341	359	.70	.738	1	.978	.757	.987
43	Ninh Thuan	.991	.896	.208	.214	.32	.332	.238	.233	.691	.973
44	Binh Phuoc	.907	.827	.099	.106	.24	.254	.153	.156	.697	.975
45	Tay Ninh	.884	.812	.170	.179	.38	.399	.497	.486	.703	.977
46	Binh Duong	.900	.857	.136	.154	.35	.385	.572	.573	.718	.981
47	Dong Nai	.952	.887	.187	.199	.29	.317	.337	.343	.715	.980
48	Binh Thuan	.986	.887	.191	.196	.35	.360	.296	.285	.707	.977
49	B. Ria - V. Tau	.926	.838	.111	.119	.42	.435	.562	.540	.743	.986
50	Long An	.915	.851	.205	.218	.28	308	.398	.403	.722	.982
51	Dong Thap	.843	.812	.121	.136	.37	.397	.415	.419	.717	.966
52	An Giang	.916	.857	.166	.175	.33	.352	.254	.258	.698	.975
53	Tien Giang	.945	.903	.155	.170	.28	.309	.312	.331	.725	.983
54	Vinh Long	.861	.844	.143	.159	.27	.303	.534	.533	.733	.985
55	Ben Tre	.928	.881	.158	.172	.30	.324	.493	.487	.700	.977
56	Kien Giang	.998	.915	.140	.147	.29	.305	.234	.231	.721	.977
57	Can Tho	.990	.945	.130	.144	.33	.358	.442	.437	.708	.972
58	Tra Vinh	.759	.727	.137	.148	.24	.261	.166	.183	.700	.972
59	Soc Trang	.928	.866	.124	.133	.23	.249	.179	.187	.696	.969
60	Bac Lieu	.920	.851	.141	.148	.29	.305	.136	.141	.697	.975
61	Ca Mau	.973	.870	.158	.161	.29	.296	.348	.328	.711	.971

4.3.2 Transformed health data for the 2002 sample

No	Province	Health Facilities								Health Status	
		Childhood immunizati-on		Hospital Beds		Medical Doctors		Pharmacist		Life Expect an-cy	Infant Mortality
		*Non Con.	** Con.	*Non Con.	** Con.	*Non Con.	** Con.	*Non Con.	** Con.		
1	Hanoi	1	1	.409	.438	.55	.612	.815	.839	.760	.991
2	Hai Phong	.997	.931	.286	.298	.65	.681	.672	.673	.739	.987
3	Vinh Phuc	.925	.919	.162	.194	.31	.371	.568	.598	.727	.984
4	Ha Tay	.954	.961	.190	.226	.38	.444	.477	.536	.706	.976
5	Bac Ninh	.854	.838	.180	.205	.48	.522	.546	.573	.720	.980
6	Hai Duong	.918	.911	.210	.235	.39	.451	.724	.747	.742	.985
7	Hung Yen	.925	.936	.192	.226	.36	.431	.381	.454	.725	.983
8	Ha Nam	.885	.896	.280	.308	.59	.651	.906	.930	.745	.986
9	Nam Dinh	.897	.885	.178	.203	.51	.570	.637	.678	.723	.981
10	Thai Binh	.999	.976	.154	.180	.56	.618	.536	.582	.726	.981
11	Ninh Binh	.880	.834	.228	.247	.44	.475	.559	.576	.720	.978
12	Ha Giang	.860	.786	.226	.236	.41	.428	.470	.460	.603	.939
13	Cao Bang	.873	.801	.320	.329	.65	.670	.494	.485	.638	.951
14	Lao Cai	.911	.827	.358	.365	.46	.473	.270	.272	.672	.965
15	Bac Kan	.854	.797	.351	.366	.77	.801	.349	.356	.677	.967
16	Lang Son	.976	.902	.230	.247	.56	.591	.584	.576	.663	.946
17	Tuyen Quang	.983	.935	.253	.275	.53	.576	.370	.390	.678	.969
18	Yen Bai	.908	.843	.290	.303	.58	.602	.580	.571	.682	.969
19	Thai Nguyen	.919	.887	.298	.322	.65	.697	.466	.487	.704	.975
20	Phu Tho	.912	.876	.195	.215	.43	.467	.415	.436	.716	.980
21	Bac Giang	.934	.909	.201	.227	.40	.455	.378	.419	.683	.970
22	Quang Ninh	.911	.846	.249	.260	.61	.634	.548	.548	.717	.980
23	Lai Chau	.918	.822	.223	.227	.45	.456	.556	.535	.650	.953
24	Son La	.937	.852	.272	.281	.42	.434	.597	.580	.673	.965
25	Hoa Binh	.942	.880	.261	.276	.42	.447	.549	.555	.682	.969
26	Thanh Hoa	.862	.786	.319	.326	.40	.412	.526	.515	.702	.972
27	Nghe An	.959	.866	.202	.209	.33	.339	.491	.477	.708	.976
28	Ha Tinh	.936	.845	.251	.256	.39	.399	.392	.387	.707	.975
29	Quang Binh	.794	.716	.216	.220	.46	.468	.703	.674	.688	.971
30	Quang Tri	.914	.826	.238	.243	.52	.532	.352	.346	.653	.958
31	Th. Thien - Hue	.884	.805	.176	.184	.45	.468	.192	.193	.709	.978
32	Da Nang	.958	.868	.389	.394	.79	.803	.483	.466	.759	.989
33	Quang Nam	.815	.753	.207	.218	.45	.472	.274	.274	.716	.978
34	Quang Ngai	.913	.831	.200	.207	.38	.394	.327	.321	.706	.972
35	Binh Dinh	.864	.782	.157	.162	.39	.400	.496	.479	.707	.974
36	Phu Yen	.904	.828	.160	.166	.51	.523	.510	.495	.672	.965

37	Khanh Hoa	.941	.863	.149	.156	.42	.437	.342	.337	.728	.976
38	KonTum	.931	.843	.368	.374	.54	.550	.383	.371	.587	.939
39	Gia Lai	.878	.808	.213	.221	.32	.338	.244	.246	.632	.957
40	Dak Lak	.927	.845	.117	.123	.37	.384	.067	.074	.667	.957
41	Lam Dong	.834	.773	.204	.212	.44	.458	.319	.315	.714	.979
42	Ho Chi Minh City	.817	.809	.311	.329	.70	.747	1	1	.760	.990
43	Ninh Thuan	.875	.798	.200	.206	.42	.434	.276	.273	.699	.975
44	Binh Phuoc	.734	.678	.110	.117	.29	.306	.153	.160	.704	.976
45	Tay Ninh	.839	.776	.174	.183	.48	.501	.499	.500	.709	.978
46	Binh Duong	.761	.736	.158	.175	.43	.469	.660	.669	.724	.983
47	Dong Nai	.866	.813	.185	.197	.32	.351	.320	.334	.721	.982
48	Binh Thuan	.850	.773	.176	.181	.41	.422	.292	.287	.713	.979
49	B. Ria - V. Tau	.922	.839	.127	.134	.56	.576	.491	.483	.748	.988
50	Long An	.931	.871	.154	.167	.37	.400	.381	.399	.727	.984
51	Dong Thap	.881	.850	.132	.145	.36	.392	.516	.524	.723	.971
52	An Giang	.938	.881	.173	.183	.35	.372	.230	.243	.705	.977
53	Tien Giang	.903	.874	.154	.168	.32	.354	.394	.416	.730	.984
54	Vinh Long	.890	.873	.140	.156	.35	.388	.455	.473	.738	.986
55	Ben Tre	.919	.881	.141	.154	.35	.380	.447	.455	.707	.978
56	Kien Giang	.964	.888	.158	.165	.30	.318	.140	.148	.727	.981
57	Can Tho	.841	.818	.132	.146	.36	.394	.487	.490	.714	.976
58	Tra Vinh	.817	.779	.134	.144	.32	.346	.190	.209	.707	.976
59	Soc Trang	.771	.732	.119	.128	.30	.326	.211	.222	.703	.974
60	Bac Lieu	.845	.785	.148	.155	.49	.507	.195	.200	.704	.977
61	Ca Mau	.764	.691	.166	.169	.34	.348	.332	.318	.717	.975

4.3.3 Transformed health data for the 2004 sample

No	Province	Health Facilities								Health Status	
		Childhood immunization		Hospital Beds		Medical Doctors		Pharmacist		Life Expectan-cy	Infant Mortality
		*NonCon.	** Con.	*Non Con.	** Con.	*Non Con.	** Con.	*NonCon.	** Con.		
1	Hanoi	1	1	.419	.450	.52	.591	1	1	.762	.992
2	Hai Phong	.987	.915	.285	.298	.68	.716	.655	.653	.742	.987
3	Vinh Phuc	.986	.968	.179	.212	.37	.433	.554	.580	.731	.984
4	Ha Tay	1	1	.194	.230	.49	.555	.336	.398	.710	.979
5	Bac Ninh	.971	.937	.191	.217	.48	.527	.334	373	.724	.982
6	Hai Duong	.969	.952	.219	.246	.41	.480	.659	.677	.745	.988
7	Hung Yen	.997	.998	.191	.226	.36	.437	.321	.386	.729	.984
8	Ha Nam	.993	.989	.279	.307	.59	.656	.744	.767	.748	.998
9	Nam Dinh	.970	.949	.178	.204	.45	.516	.637	.667	.727	.983
10	Thai Binh	1	.975	.186	.212	.71	.767	.521	.560	.729	.984
11	Ninh Binh	.989	.928	.224	.240	.43	.464	.483	.498	.724	.981
12	Ha Giang	.966	.873	.236	.247	.46	.479	.469	.458	.613	.942
13	Cao Bang	.908	.825	.336	.346	.73	.754	.492	.481	.647	.954
14	Lao Cai	.989	890	.457	.464	.52	.532	.336	.332	.679	.968
15	Bac Kan	.920	.850	.363	.378	.86	.896	.270	.279	.684	.970
16	Lang Son	.990	.908	.239	.254	.84	.875	.615	.602	.670	.948
17	Tuyen Quang	.997	.944	.262	.285	.52	.570	.362	.378	.684	.970
18	Yen Bai	.993	.913	.292	.306	.61	.633	.567	.557	.687	.971
19	Thai Nguyen	.986	.941	.308	.333	.67	.726	.365	.392	.708	.978
20	Phu Tho	.981	.932	.215	.235	.44	.479	.342	.363	.720	.982
21	Bac Giang	.991	.955	.214	.241	.60	.661	.397	.432	.689	.972
22	Quang Ninh	.969	.890	.256	.267	.66	.690	.525	.524	.721	.982
23	Lai Chau	.945	.839	.256	.261	.25	.256	.507	.487	.649	.957
24	Son La	.985	.888	.274	.282	.37	.385	.555	.539	.680	.968
25	Hoa Binh	.980	.910	.229	.243	.44	.467	.548	.547	.688	.971
26	Thanh Hoa	.984	.888	.229	.235	.43	.442	.469	.458	.706	.975
27	Nghe An	.986	.884	.201	.207	.34	.350	.360	.351	.712	.978
28	Ha Tinh	.984	.882	.274	.279	.41	.420	.404	.392	.711	.978
29	Quang Binh	.974	.868	.222	.227	.51	.519	.469	.451	.693	.973
30	Quang Tri	.962	.864	.252	.257	.54	.553	.341	.331	.660	.961
31	Th. Thien - Hue	.981	.884	.185	.194	.48	.499	.232	.233	.713	.979
32	Da Nang	.954	.860	.380	.385	.90	.913	.733	.701	.761	.991
33	Quang Nam	.974	.887	.226	.238	.46	.483	.296	.296	.720	.981
34	Quang Ngai	.980	.885	.189	.196	.42	.435	.341	.334	.710	.975
35	Binh Dinh	.985	.882	.169	.173	.43	.441	.569	.547	.711	.977
36	Phu Yen	.971	.881	.159	.165	.50	.516	.495	.479	.678	.968
37	Khanh Hoa	.981	.893	.165	.172	.48	.498	.414	.403	.732	.978
38	KonTum	.951	.855	.366	.372	.51	.522	.300	.292	.597	.942
39	Gia Lai	.908	.829	.213	.221	.45	.468	.173	.178	.642	.960
40	Dak Lak	.959	.869	.134	.140	.43	.446	.058	.065	.675	.967
41	Lam Dong	.936	.860	.198	.206	.46	.481	.325	.320	.718	.981

42	Ho Chi Minh City	.878	.866	.334	.351	.75	.797	1	.994	.762	.992
43	Ninh Thuan	.987	.892	.201	.207	.45	.466	.288	.285	.704	.977
44	Binh Phuoc	.939	.856	.135	.141	.41	.427	.128	.133	.708	.978
45	Tay Ninh	.995	.911	.080	.090	.49	.514	.239	.251	.713	.979
46	Binh Duong	.990	.940	.175	.192	.44	.483	.600	.604	.728	.984
47	Dong Nai	.975	.908	.189	.202	.32	.353	.345	.358	.725	.983
48	Binh Thuan	.923	.832	.176	.181	.49	.502	.326	.319	.717	.981
49	B. Ria - V. Tau	.984	.888	.129	.137	.45	.467	.590	.577	.750	.989
50	Long An	.935	.873	.146	.157	.40	.432	.450	.458	.731	.985
51	Dong Thap	.981	.935	.149	.162	.43	.467	.518	.526	.726	.974
52	An Giang	.969	.905	.176	.186	.36	.391	.184	.198	.709	.978
53	Tien Giang	.986	.944	.166	.180	.36	.401	.351	.377	.733	.985
54	Vinh Long	.967	.943	.142	.158	.39	.438	.565	.574	.741	.987
55	Ben Tre	.973	.928	.156	.169	.49	.524	.424	.433	.711	.979
56	Kien Giang	.949	.871	.180	.188	.46	.482	.104	.113	.730	.984
57	Can Tho	.991	.949	.129	.144	.54	.580	.483	.485	.724	.983
58	Tra Vinh	.937	.885	.135	.145	.38	.413	.187	.208	.711	.978
59	Soc Trang	.934	.876	.132	.142	.32	.351	.215	.227	.707	.976
60	Bac Lieu	.967	.889	.148	.157	.53	.553	.216	.220	.708	.978
61	Ca Mau	.849	.760	.194	.197	.44	.450	.391	.374	.721	.977

Note: *(*non-con.*) implies non-contiguity; **(*con.*) stands for contiguity

Appendices to Chapter 5

5.1 Transformed education data*

5.1.1 Transformed education data for the 1998 sample

No	Province	Education Facility indices						Enrolment indices		
		Primary education institution	Primary education teacher	Lower secondary education institution	Lower secondary education teacher	Upper secondary education institution	Upper secondary education teacher	Primary enrolment	Lower secondary enrolment	Upper secondary enrolment
1	Hanoi	0.1076	0.3223	0.1141	0.4622	0.0815	0.4141	0.958	0.837	0.600
2	Hai Phong	0.1082	0.3108	0.1295	0.3663	0.0967	0.4792	0.933	0.802	0.494
3	Vinh Phuc	0.1051	0.2959	0.1279	0.3123	0.1395	0.2344	0.946	0.792	0.379
4	Ha Tay	0.1149	0.3313	0.1571	0.3989	0.0767	0.3093	0.902	0.731	0.462
5	Bac Ninh	0.0980	0.2970	0.1458	0.3162	0.1032	0.3409	0.908	0.836	0.427
6	Hai Duong	0.1326	0.3164	0.1715	0.3125	0.0810	0.2978	0.945	0.782	0.443
7	Hung Yen	0.1169	0.2856	0.1752	0.3167	0.0812	0.2759	0.909	0.780	0.378
8	Ha Nam	0.1385	0.2931	0.1679	0.3226	0.0921	0.2894	0.918	0.776	0.360
9	Nam Dinh	0.1266	0.2902	0.1415	0.3259	0.0916	0.3694	0.950	0.829	0.335
10	Thai Binh	0.1532	0.3138	0.1889	0.4113	0.0734	0.3211	0.916	0.757	0.417
11	Ninh Binh	0.1140	0.3081	0.1689	0.3382	0.1006	0.2963	0.892	0.669	0.354
12	Ha Giang	0.1630	0.5598	0.0893	0.5569	0.3221	0.4669	0.579	0.332	0.095
13	Cao Bang	0.2601	0.4258	0.1446	0.4998	0.2987	0.5206	0.668	0.451	0.127
14	Lao Cai	0.1701	0.3910	0.1483	0.4928	0.2787	0.6052	0.656	0.331	0.141
15	Bac Kan	0.1337	0.4219	0.2065	0.4491	0.1781	0.4964	0.634	0.353	0.324
16	Lang Son	0.0912	0.3965	0.0604	0.4649	0.1861	0.4362	0.809	0.545	0.158
17	Tuyen Quang	0.0891	0.4028	0.1324	0.4699	0.1831	0.3273	0.811	0.538	0.221
18	Yen Bai	0.1377	0.3726	0.1276	0.4858	0.1392	0.4489	0.802	0.586	0.284
19	Thai Nguyen	0.0726	0.3666	0.0796	0.4487	0.0869	0.4288	0.948	0.764	0.355
20	Phu Tho	0.1486	0.3411	0.1759	0.3766	0.1158	0.2852	0.921	0.711	0.386
21	Bac Giang	0.0904	0.3348	0.1435	0.3441	0.1281	0.3308	0.870	0.740	0.302
22	Quang Ninh	0.1025	0.3442	0.0831	0.4281	0.1116	0.3033	0.877	0.732	0.363
23	Lai Chau	0.1678	0.4349	0.0448	0.4499	0.2991	0.5846	0.632	0.258	0.109
24	Son La	0.0947	0.3999	0.0894	0.3808	0.2447	0.4787	0.699	0.431	0.132
25	Hoa Binh	0.1188	0.4399	0.1621	0.4992	0.1839	0.4562	0.861	0.697	0.335
26	Thanh Hoa	0.1059	0.2932	0.1712	0.3147	0.0983	0.3553	0.928	0.659	0.347
27	Nghe An	0.1244	0.3434	0.1532	0.3286	0.0936	0.3188	0.908	0.699	0.386
28	Ha Tinh	0.1508	0.2927	0.1999	0.3511	0.0927	0.3331	0.851	0.511	0.306
29	Quang Binh	0.1627	0.3296	0.1925	0.4212	0.1368	0.3715	0.830	0.513	0.216
30	Quang Tri	0.1904	0.3393	0.1783	0.3404	0.1518	0.5037	0.700	0.509	0.306
31	Th.Thien-Hue	0.1516	0.3173	0.0919	0.3584	0.1121	0.3712	0.933	0.669	0.306
32	Da Nang	0.0894	0.3158	0.0699	0.4164	0.0719	0.4349	0.876	0.684	0.342
33	Quang Nam	0.1119	0.3436	0.1054	0.3583	0.0921	0.3147	0.853	0.709	0.362

34	Quang Ngai	0.1306	0.3379	0.1243	0.3693	0.1166	0.3584	0.838	0.540	0.402
35	Binh Dinh	0.1198	0.2937	0.1055	0.3149	0.1214	0.3273	0.846	0.621	0.362
36	Phu Yen	0.1062	0.3684	0.1025	0.4281	0.1309	0.4260	0.822	0.609	0.280
37	Khanh Hoa	0.1108	0.3064	0.0815	0.4142	0.1297	0.4263	0.846	0.686	0.349
38	KonTum	0.1171	0.3775	0.0484	0.3921	0.2461	0.5414	0.751	0.592	0.187
39	Gia Lai	0.0996	0.3102	0.0942	0.3194	0.1614	0.3410	0.813	0.665	0.239
40	Dak Lak	0.0917	0.3152	0.0848	0.3185	0.1455	0.3517	0.799	0.634	0.274
41	Lam Dong	0.1329	0.3116	0.0862	0.3679	0.1586	0.4289	0.829	0.655	0.341
42	Ho Chi Minh	0.0966	0.3027	0.0618	0.3527	0.0739	0.3835	0.859	0.762	0.433
43	Ninh Thuan	0.1509	0.3307	0.0976	0.4369	0.1227	0.3390	0.702	0.504	0.252
44	Binh Phuoc	0.1037	0.2929	0.1041	0.1749	0.1967	0.2803	0.766	0.525	0.305
45	Tay Ninh	0.1797	0.3481	0.1028	0.3438	0.1673	0.3328	0.803	0.664	0.215
46	Binh Duong	0.1199	0.3278	0.0783	0.3113	0.1588	0.2969	0.808	0.505	0.329
47	Dong Nai	0.0838	0.2433	0.0636	0.2340	0.0864	0.2182	0.798	0.644	0.376
48	Binh Thuan	0.1142	0.2843	0.0869	0.3187	0.1076	0.2519	0.730	0.545	0.265
49	B. Ria-V. Tau	0.1155	0.3059	0.0842	0.2997	0.1015	0.3953	0.803	0.679	0.410
50	Long An	0.1189	0.3206	0.0702	0.3747	0.1562	0.3180	0.778	0.593	0.229
51	Dong Thap	0.1319	0.3068	0.1155	0.3268	0.1375	0.4549	0.728	0.458	0.183
52	An Giang	0.1252	0.2443	0.0749	0.3059	0.1379	0.3043	0.655	0.441	0.167
53	Tien Giang	0.1053	0.3167	0.0907	0.3557	0.1057	0.3484	0.805	0.635	0.277
54	Vinh Long	0.1679	0.3276	0.1079	0.3484	0.1316	0.2773	0.756	0.547	0.266
55	Ben Tre	0.0869	0.3129	0.0741	0.3149	0.1188	0.3239	0.764	0.628	0.268
56	Kien Giang	0.0702	0.2953	0.0807	0.3194	0.1706	0.3559	0.671	0.410	0.132
57	Can Tho	0.0696	0.2733	0.0355	0.2887	0.1479	0.2854	0.708	0.474	0.182
58	Tra Vinh	0.1158	0.2852	0.1082	0.2899	0.0976	0.2223	0.713	0.460	0.103
59	Soc Trang	0.1492	0.3017	0.1479	0.2683	0.1120	0.2577	0.607	0.372	0.115
60	Bac Lieu	0.1053	0.2906	0.1474	0.3139	0.1340	0.3470	0.704	0.383	0.137
61	Ca Mau	0.0909	0.3073	0.1001	0.2472	0.1691	0.2251	0.730	0.614	0.137

5.1.2 Transformed education data for the 2002 sample

No	Province	Education Facility Indices						Enrolment Indices		
		Primary education institution	Primary education teacher	Lower secondary education institution	Lower secondary education teacher	Upper secondary education institution	Upper secondary education teacher	Primary enrolment	Lower secondary enrolment	Upper secondary enrolment
1	Hanoi	0.1237	0.3944	0.1237	0.5259	0.0951	0.4339	0.954	0.867	0.558
2	Hai Phong	0.1453	0.4053	0.1336	0.4819	0.0811	0.5093	0.919	0.888	0.449
3	Vinh Phuc	0.1663	0.4202	0.1529	0.4708	0.0756	0.5393	0.949	0.849	0.413
4	Ha Tay	0.1530	0.4838	0.1570	0.5373	0.0564	0.3096	0.923	0.835	0.535
5	Bac Ninh	0.1506	0.4144	0.1473	0.4082	0.0679	0.3832	0.929	0.904	0.478
6	Hai Duong	0.1909	0.4302	0.1792	0.4188	0.0673	0.3229	0.900	0.937	0.431
7	Hung Yen	0.1630	0.4173	0.1643	0.3887	0.0688	0.3271	0.879	0.941	0.436
8	Ha Nam	0.1776	0.3929	0.1702	0.4562	0.0889	0.3702	0.930	0.893	0.404
9	Nam Dinh	0.1592	0.3591	0.1493	0.4366	0.00697	0.3732	0.955	0.886	0.399
10	Thai Binh	0.1977	0.4932	0.1943	0.4998	0.0690	0.3316	0.891	0.929	0.456
11	Ninh Binh	0.1709	0.4189	0.1603	0.4403	0.0697	0.3260	0.947	0.881	0.346
12	Ha Giang	0.1510	0.6084	0.0893	0.4868	0.1197	0.7501	0.945	0.419	0.115
13	Cao Bang	0.2712	0.5540	0.1324	0.4317	0.1435	0.3569	0.913	0.608	0.220
14	Lao Cai	0.2526	0.4803	0.1695	0.4965	0.1633	0.4145	0.936	0.574	0.159
15	Bac Kan	0.3059	0.5874	0.2254	0.5504	0.0958	0.3268	0.820	0.551	0.315
16	Lang Son	0.1517	0.5122	0.1024	0.5059	0.0907	0.3262	0.945	0.728	0.225
17	Tuyen Quang	0.1349	0.5432	0.1397	0.4522	0.0722	0.2906	0.957	0.883	0.384
18	Yen Bai	0.1829	0.5139	0.1361	0.5013	0.0975	0.4255	0.801	0.732	0.311
19	Thai Nguyen	0.2161	0.5272	0.1774	0.5109	0.0619	0.4307	0.948	0.764	0.339
20	Phu Tho	0.2261	0.4879	0.2013	0.4755	0.0959	0.2857	0.956	0.861	0.422
21	Bac Giang	0.1498	0.4901	0.1328	0.3846	0.0682	0.3279	0.916	0.920	0.349
22	Quang Ninh	0.1347	0.4653	0.1229	0.5110	0.0827	0.3811	0.996	0.810	0.422
23	Lai Chau	0.1496	0.5310	0.1355	0.5229	0.1169	0.4816	0.910	0.386	0.130
24	Son La	0.1061	0.5177	0.1304	0.4887	0.0941	0.3674	0.929	0.455	0.173
25	Hoa Binh	0.2140	0.6678	0.2229	0.5962	0.1434	0.4111	0.951	0.763	0.328
26	Thanh Hoa	0.1734	0.4098	0.1767	0.3628	0.0685	0.3339	0.988	0.878	0.385
27	Nghe An	0.1785	0.4446	0.1375	0.4006	0.0675	0.3848	0.913	0.977	0.499
28	Ha Tinh	0.1911	0.3902	0.1453	0.4285	0.0601	0.3329	0.908	0.972	0.510
29	Quang Binh	0.2369	0.3698	0.1659	0.4511	0.0682	0.3397	0.876	0.822	0.358
30	Quang Tri	0.2046	0.4032	0.1614	0.4235	0.0889	0.5086	0.994	0.841	0.493
31	Th.Thien-Hue	0.1619	0.3222	0.0917	0.3806	0.0593	0.4185	0.941	0.793	0.378
32	Da Nang	0.1127	0.3639	0.0775	0.4457	0.0635	0.4316	0.905	0.931	0.548
33	Quang Nam	0.1496	0.3924	0.1140	0.3971	0.0699	0.3163	0.918	0.921	0.494
34	Quang Ngai	0.1507	0.3512	0.1153	0.3829	0.0615	0.4028	0.947	0.869	0.409
35	Binh Dinh	0.1250	0.3256	0.0789	0.3364	0.0499	0.3348	0.924	0.880	0.358
36	Phu Yen	0.1467	0.4771	0.1154	0.4899	0.0651	0.4928	0.921	0.857	0.386
37	Khanh Hoa	0.1228	0.3456	0.0827	0.3950	0.0428	0.4178	0.909	0.828	0.293
38	KonTum	0.1267	0.4381	0.1111	0.5388	0.1273	0.4490	0.758	0.807	0.210
39	Gia Lai	0.0960	0.3799	0.1015	0.4035	0.0866	0.3735	0.795	0.661	0.262

40	Dak Lak	0.1188	0.3483	0.1001	0.3520	0.0710	0.3163	0.801	0.721	0.270
41	Lam Dong	0.1713	0.3603	0.0957	0.4239	0.0864	0.4563	0.846	0.832	0.306
42	Ho Chi Minh	0.1020	0.3189	0.0687	0.4079	0.0719	0.4226	0.996	0.762	0.433
43	Ninh Thuan	0.1619	0.3691	0.0795	0.3892	0.0829	0.3729	0.913	0.621	0.202
44	Binh Phuoc	0.1168	0.3870	0.1076	0.3629	0.0929	0.2877	0.789	0.683	0.272
45	Tay Ninh	0.2659	0.4386	0.1187	0.4048	0.1249	0.3359	0.973	0.607	0.247
46	Binh Duong	0.1514	0.4565	0.0813	0.4271	0.1267	0.3441	0.834	0.761	0.224
47	Dong Nai	0.1132	0.3412	0.0768	0.3308	0.0549	0.2873	0.984	0.821	0.380
48	Binh Thuan	0.1440	0.3413	0.0864	0.3281	0.0756	0.3561	0.980	0.700	0.263
49	B. Ria-V. Tau	0.1293	0.3609	0.0756	0.3699	0.0827	0.3500	0.816	0.781	0.336
50	Long An	0.1591	0.4081	0.0746	0.3867	0.0929	0.3124	0.942	0.728	0.210
51	Dong Thap	0.1617	0.4004	0.0807	0.3742	0.0883	0.4005	0.986	0.641	0.202
52	An Giang	0.1812	0.3421	0.0886	0.3390	0.0754	0.3253	0.920	0.545	0.177
53	Tien Giang	0.1467	0.3555	0.0969	0.4095	0.0759	0.3313	0.805	0.635	0.277
54	Vinh Long	0.2657	0.4079	0.1038	0.3811	0.0826	0.2893	0.957	0.789	0.357
55	Ben Tre	0.1461	0.4086	0.1198	0.4171	0.0973	0.2868	0.999	0.765	0.350
56	Kien Giang	0.1050	0.3877	0.0878	0.3727	0.0927	0.3251	0.934	0.513	0.153
57	Can Tho	0.1659	0.4055	0.0963	0.4939	0.0664	0.2108	0.939	0.597	0.143
58	Tra Vinh	0.1892	0.4287	0.1076	0.3956	0.0758	0.4903	0.879	0.606	0.168
59	Soc Trang	0.1480	0.4083	0.0874	0.3418	0.0959	0.3503	0.937	0.488	0.111
60	Bac Lieu	0.1413	0.3981	0.1291	0.3752	0.0753	0.3099	0.745	0.540	0.141
61	Ca Mau	0.1535	0.4303	0.1107	0.3994	0.0904	0.2893	0.948	0.505	0.116

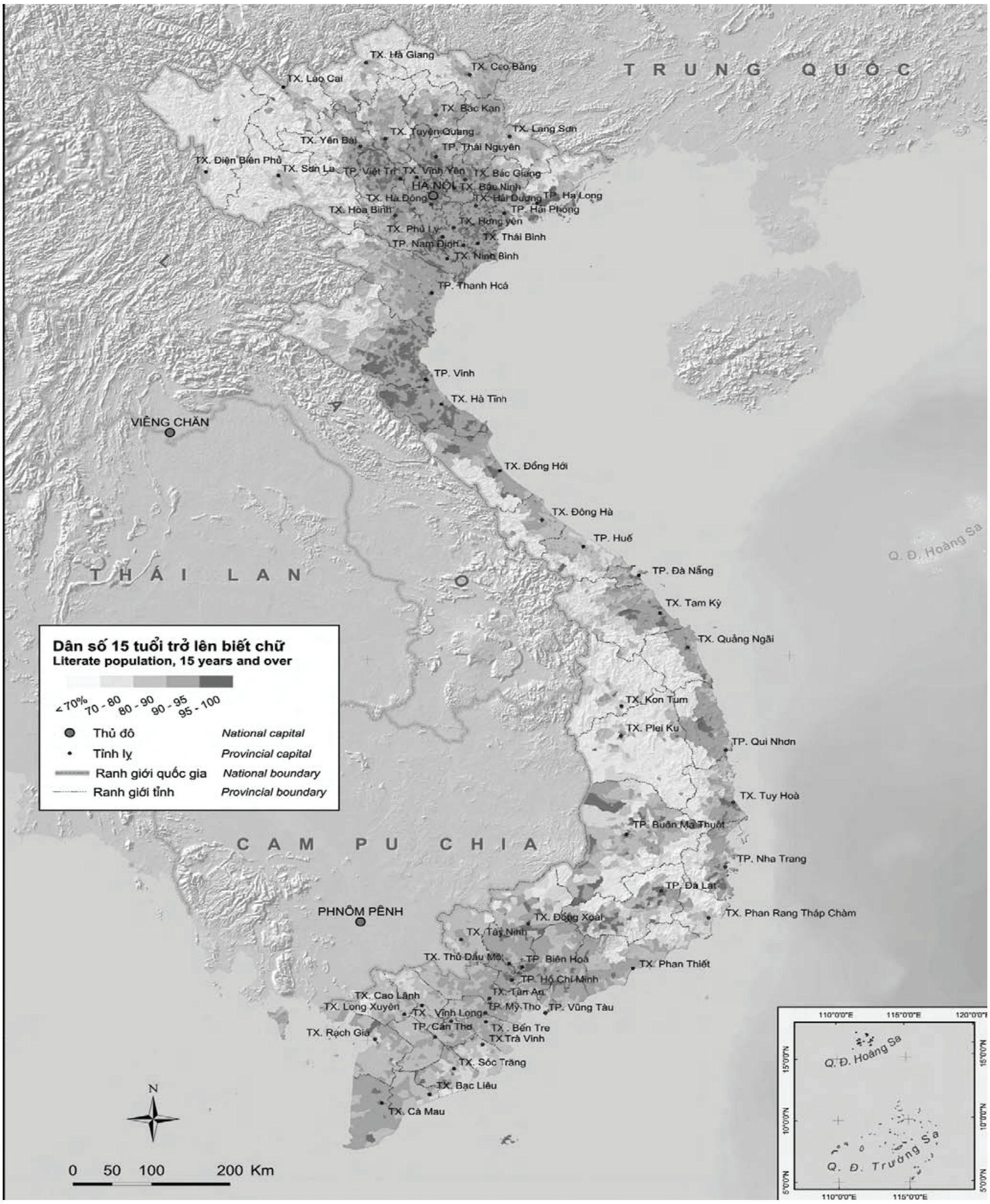
5.1.3 Transformed education data for the 2004 sample

No	Province	Education Facility indices						Enrolment indices		
		Primary education institution	Primary education teacher	Lower secondary education institution	Lower second-dary education teacher	Upper secondary education institution	Upper second-dary education teacher	Primary enroll-ment	Lower second-dary enroll-ment	Upper second-dary enroll-ment
1	Hanoi	0.1308	0.4256	0.1213	0.4754	0.0929	0.4814	0.967	0.951	0.627
2	Hai Phong	0.1684	0.4748	0.1382	0.4450	0.0767	0.4822	0.916	0.891	0.389
3	Vinh Phuc	0.1920	0.4638	0.1519	0.4562	0.0804	0.4039	0.921	0.867	0.514
4	Ha Tay	0.1704	0.5422	0.1639	0.5557	0.0529	0.2959	0.877	0.836	0.556
5	Bac Ninh	0.1688	0.4363	0.1488	0.4374	0.0614	0.4234	0.926	0.920	0.599
6	Hai Duong	0.2160	0.4489	0.1928	0.4126	0.0666	0.3204	0.920	0.899	0.473
7	Hung Yen	0.1896	0.4624	0.1695	0.4186	0.1127	0.3376	0.893	0.935	0.560
8	Ha Nam	0.2118	0.4236	0.1732	0.3991	0.0893	0.4203	0.925	0.931	0.456
9	Nam Dinh	0.1834	0.3945	0.1500	0.3885	0.0713	0.4087	0.924	0.934	0.439
10	Thai Binh	0.2227	0.5234	0.1986	0.4954	0.0629	0.3413	0.890	0.928	0.559
11	Ninh Binh	0.2130	0.4945	0.1641	0.4128	0.0707	0.3525	0.914	0.930	0.482
12	Ha Giang	0.1488	0.6733	0.0956	0.8353	0.1072	0.4181	0.949	0.341	0.122
13	Cao Bang	0.3010	0.6724	0.1442	0.8199	0.0648	0.2985	0.860	0.633	0.310
14	Lao Cai	0.2613	0.5821	0.2493	0.7366	0.1509	0.4454	0.820	0.468	0.136
15	Bac Kan	0.3648	0.5979	0.2711	0.6695	0.0968	0.3075	0.772	0.493	0.307
16	Lang Son	0.2059	0.6015	0.1331	0.6534	0.0705	0.3071	0.894	0.726	0.288
17	Tuyen Quang	0.2086	0.6435	0.1576	0.6064	0.0334	0.3639	0.908	0.940	0.614
18	Yen Bai	0.1984	0.5475	0.1595	0.6435	0.0784	0.4022	0.893	0.761	0.359
19	Thai Nguyen	0.2564	0.5921	0.1964	0.5682	0.0601	0.4251	0.906	0.770	0.324
20	Phu Tho	0.2704	0.6067	0.2079	0.5534	0.0868	0.4053	0.933	0.902	0.531
21	Bac Giang	0.1792	0.5412	0.1459	0.5149	0.0649	0.3763	0.925	0.910	0.435
22	Quang Ninh	0.1523	0.5302	0.1403	0.6007	0.0812	0.4026	0.991	0.797	0.353
23	Lai Chau	0.1399	0.6119	0.1841	0.7572	0.1478	0.4499	0.979	0.460	0.168
24	Son La	0.1353	0.5594	0.1736	0.9236	0.0774	0.3516	0.981	0.563	0.171
25	Hoa Binh	0.2871	0.7704	0.2591	0.7549	0.1174	0.3940	0.871	0.866	0.297
26	Thanh Hoa	0.2251	0.4972	0.1779	0.4386	0.0639	0.3549	0.969	0.347	0.436
27	Nghe An	0.2167	0.5147	0.1406	0.5054	0.0635	0.3925	0.969	0.994	0.461
28	Ha Tinh	0.2297	0.4453	0.1396	0.4306	0.0606	0.3632	0.989	0.974	0.635
29	Quang Binh	0.2628	0.4212	0.1631	0.4412	0.0707	0.3662	0.965	0.951	0.497
30	Quang Tri	0.2099	0.4795	0.1495	0.5567	0.0954	0.4947	0.937	0.884	0.478
31	Th.Thien-Hue	0.1791	0.3681	0.0926	0.4812	0.0653	0.5570	0.972	0.865	0.442
32	Da Nang	0.1373	0.4191	0.0766	0.4627	0.0542	0.4159	0.951	0.922	0.625
33	Quang Nam	0.1775	0.4429	0.1193	0.4682	0.0671	0.3134	0.906	0.876	0.307
34	Quang Ngai	0.1763	0.4066	0.1108	0.4468	0.0623	0.3936	0.931	0.819	0.440
35	Binh Dinh	0.1563	0.3849	0.0734	0.4252	0.0540	0.3023	0.907	0.827	0.498
36	Phu Yen	0.1766	0.5416	0.1259	0.6622	0.0802	0.5901	0.855	0.847	0.333
37	Khanh Hoa	0.1505	0.3944	0.0784	0.4812	0.0458	0.3780	0.942	0.778	0.367
38	KonTum	0.1709	0.5220	0.1559	0.9230	0.1138	0.4965	0.971	0.962	0.289
39	Gia Lai	0.1119	0.3877	0.1256	0.6939	0.0831	0.3669	0.758	0.655	0.346
40	Dak Lak	0.1447	0.3971	0.1076	0.5798	0.0456	0.3241	0.977	0.768	0.384
41	Lam Dong	0.1897	0.3997	0.0983	0.5340	0.0410	0.4403	0.841	0.888	0.362

42	Ho Chi Minh	0.1073	0.3576	0.0715	0.4727	0.0448	0.4279	0.906	0.861	0.516
43	Ninh Thuan	0.1952	0.3998	0.0812	0.5935	0.0926	0.4034	0.950	0.668	0.318
44	Binh Phuoc	0.1517	0.4597	0.1239	0.6651	0.0839	0.3563	0.929	0.684	0.339
45	Tay Ninh	0.3139	0.5032	0.1312	0.6462	0.1063	0.3280	0.863	0.711	0.272
46	Binh Duong	0.1675	0.4844	0.0796	0.6124	0.1097	0.3829	0.985	0.758	0.411
47	Dong Nai	0.1431	0.3812	0.0792	0.4623	0.0641	0.3215	0.897	0.864	0.367
48	Binh Thuan	0.1834	0.3981	0.0878	0.5486	0.0556	0.3496	0.987	0.751	0.402
49	B. Ria-V. Tau	0.0339	0.0866	0.0187	0.1144	0.0135	0.0857	0.981	0.790	0.444
50	Long An	0.1979	0.4364	0.1034	0.5378	0.0742	0.3248	0.975	0.778	0.370
51	Dong Thap	0.1985	0.4683	0.0843	0.5677	0.0589	0.4608	0.826	0.478	0.248
52	An Giang	0.2011	0.3913	0.1016	0.6066	0.1163	0.3725	0.789	0.611	0.268
53	Tien Giang	0.1653	0.3825	0.1036	0.4752	0.0744	0.3589	0.953	0.812	0.371
54	Vinh Long	0.3032	0.5072	0.1116	0.5345	0.0781	0.3781	0.907	0.859	0.436
55	Ben Tre	0.1889	0.4585	0.1438	0.5274	0.0920	0.3561	0.942	0.798	0.480
56	Kien Giang	0.1350	0.4414	0.0912	0.6721	0.0849	0.3892	0.879	0.629	0.225
57	Can Tho	0.3678	0.4653	0.1432	0.6254	0.1686	0.3208	0.988	0.592	0.256
58	Tra Vinh	0.2539	0.5371	0.1136	0.6255	0.0762	0.4815	0.942	0.705	0.305
59	Soc Trang	0.1817	0.4804	0.1056	0.8048	0.0899	0.4688	0.951	0.512	0.219
60	Bac Lieu	0.1758	0.4662	0.1232	0.7366	0.0806	0.2931	0.758	0.524	0.139
61	Ca Mau	0.2077	0.5551	0.1148	0.8455	0.0870	0.3405	0.771	0.561	0.130

** The data are rounded.*

Appendix 5.2
Spatial Distribution of the literate population 15 years and over



Appendices to Chapter 6

6.1 GDP per capita by provinces in selected years*

Code	Provinces\years	1990	1995	2000	2006
101	Hanoi	1097	5964	11504	28269
102	Hai Phong	851	3817	6018	14187
103	Vinh Phuc	261	1510	3545	10992
104	Ha Tay	410	1898	3157	7028
105	Bac Ninh	382	1855	3548	10092
106	Hai Duong	521	2404	3712	9017
107	Hung Yen	405	1969	3802	8367
108	Ha Nam	403	1870	2995	6188
109	Nam Dinh	483	1992	2891	6216
110	Thai Binh	512	2229	3183	6751
111	Ninh Binh	346	1610	2695	6396
201	Ha Giang	149	814	1721	3545
202	Cao Bang	267	1432	2736	5243
203	Lao Cai	414	1563	2141	6168
204	Bac Kan	252	1119	1753	3970
205	Lang Son	387	1819	3077	6410
206	Tuyen Quang	346	1551	2406	5502
207	Yen Bai	291	1387	2418	5036
208	Thai Nguyen	449	1862	2860	6911
209	Phu Tho	418	1856	2999	6060
210	Bac Giang	389	1603	2341	5113
211	Quang Ninh	634	3040	5338	13857
301	Lai Chau	394	1510	2315	4682
302	Son La	290	1279	2028	4741
303	Hoa Binh	389	1627	2292	9706
401	Thanh Hoa	462	1908	2851	6016
402	Nghe An	433	1855	2748	6405
403	Ha Tinh	429	1801	2668	4676
404	Quang Binh	362	1611	2765	6461
405	Quang Tri	377	1738	2890	6581
406	Thua Thien-Hue	431	2054	3254	7406
501	Da Nang	685	3673	7031	17589
502	Quang Nam	324	1668	3054	7206
503	Quang Ngai	329	1571	2690	6133
504	Binh Dinh	436	1948	3100	7851
505	Phu Yen	325	1719	3066	7198
506	Khanh Hoa	675	3444	6022	13947
601	Kon Tum	310	1549	2856	6288
602	Gia Lai	408	1743	2599	5529
603	Dak Lak	765	2566	2773	6304
604	Lam Dong	468	2006	2812	7831
701	Ho Chi Minh City	1849	8806	14516	31283
702	Ninh Thuan	575	2228	2959	5447
703	Binh Phuoc	453	1845	2642	7629
704	Tay Ninh	471	2372	4141	11959
705	Binh Duong	-	-	8224	18525
706	Dong Nai	727	3660	6676	16295
707	Binh Thuan	373	1742	2909	8782

708	Ba Ria-Vung Tau	4979	21876	51063	166995
801	Long An	709	3047	4499	9779
802	Dong Thap	509	2215	3434	7264
803	An Giang	559	2799	4560	9512
804	Tien Giang	587	2676	4261	8611
805	Vinh Long	573	2720	4247	9073
806	Ben Tre	516	2457	4149	8245
807	Kien Giang	729	3190	4750	11261
808	Can Tho	670	2902	4675	12191
809	Tra Vinh	536	2387	4272	9252
810	Soc Trang	578	2570	4227	8093
811	Bac Lieu	524	2525	4212	11484
812	Ca Mau	784	3346	5234	10255
National average		*568*	*2587*	*4628*	*11571*

Source: Statistical Year Book, Various Years

** The figure is measured in thousand Vietnamese Dong. The figures are rounded.*

Annex to statement
Name: Nguyen Huy Hoang
PhD student, Mansholt Graduate School of Social Sciences (MG3S)
Completed Training and Supervision Plan

Description	Institute / Department	Year	ECTS*
Courses:			34
Mansholt Introduction course	Mansholt Graduate School of Social Sciences (MG3S)	2004	1.5
Advanced econometrics	MG3S	2005	6
Time planning and project management	Wageningen Graduate Schools (WGS)	2006	1,5
Research Methodology: Designing and Conducting a PhD research project	MG3S	2004	2.8
The Empirics of economic Growth	NAKE	2005	3
Economic Organization Theory	NAKE	2005	3
Techniques for Writing and Presenting a Scientific Paper	WGS	2006	1.2
Development Economics	NAKE	2006	3
Introduction to Environment Economics modelling	SENSE Research School	2006	6
Microeconomics	ECH 21806	2005	6
Presentations at conferences and workshops:			3
Various presentations about Regional Integration and Economic Cooperation and Development in ASEAN at the Institute for Southeast Asian Studies, Vietnam Academy of Social Sciences, Hanoi, Vietnam			1
Public Economics Theory 2006 Hanoi International Conference		2006	1
Policies to ensure rapid and sustainable growth in the current period 2009 Hanoi International Conference		2009	1
Total (minimum 30 ECTS)			37