ANALYSING THE ROLE OF INFRASTRUCTURE IN VIETNAM

By

HUONG MAI THI TRUONG

A thesis submitted to the University of Birmingham

For the degree of

DOCTOR OF PHILOSOPHY

Department of Economics

Birmingham Business School

College of Social Sciences

The University of Birmingham

September 2012

ABSTRACT

The thesis contains three empirical studies into economic effects of infrastructure, focusing on different aspects of the Vietnamese economy during the period 2000-2007, namely, economic growth, private sector employment, and private sector location choices. In addition, the empirical studies explicitly take into account the potential existence of cross-province transport infrastructure spillovers. By estimating empirical models specified in accordance with the relevant literature and the context of the Vietnamese economy, and subject to the availability of data, the thesis obtains the following main findings. Firstly, there is evidence of a positive impact of transport infrastructure on economic growth, on private sector employment, and on private sector location choices. Secondly, sectoral differences in the impact of transport infrastructure are identified in the location choice analysis. Thirdly, the cross-province transport infrastructure spillovers cannot be seen in the growth analysis; whereas, there is some evidence of a negative spillover effect on private sector employment. In the location choice analysis, the evidence of spillovers varies dramatically. Finally, higher-education infrastructure, which is assumed to be associated with the capacity of qualified labour supply, is positively related to both private sector employment and location choices.

ACKNOWLEDGMENTS

My gratitude goes to my supervisors Somnath Sen and Nicholas Horsewood for their advice and guidance throughout the three years of my study at the University of Birmingham.

This research has been impossible without the support of the Ministry of Education and Training, the Ministry of Transport, the General Statistics Office, and the Central Institute for Economic Management, which provided data.

All errors are my own.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ADB	Asian Development Bank
EPZ	Export Processing Zone
FDI	Foreign Direct Investment
FIE	Foreign-Invested Enterprise
FGLS	Feasible Generalised Least Squares
GSO	General Statistics Office of Vietnam
GDP	Gross Domestic Product
GMM	Generalised Method of Moments
IZ	Industrial Zone
NSE	Non-State Enterprise
OLS	Ordinary Least Squares
SOE	State Owned Enterprise

CHAPTER 1

INTRODUCTION

An adequate infrastructure endowment is widely viewed as one of the most essential prerequisites for development. From a policy perspective, infrastructure investment is considered a traditional approach to development. In Vietnam, especially since the introduction of the Renovation policy, known as *Đổi Mới*, in 1986, a great deal of emphasis has been put on infrastructure investment. With respect to economic literature, the beneficial impact of infrastructure has been investigated in several contexts. However, the literature related to developing countries, in particular Vietnam, is still limited. Therefore, this thesis presents an empirical research on the impacts of infrastructure on three aspects of the Vietnamese economy, namely, economic growth, private sector employment and private sector location choices, during the period 2000-2007.

Chapter 1 begins with a presentation of the background and motivation of the thesis, including a review of infrastructure concept, an overview of the literature on the economic role of infrastructure, and a brief description on the performance of the Vietnamese economy during the period 2000-2007. Terminologies, research questions, methodology and an outline of the thesis are presented in the subsequent sections.

1.1 BACKGROUND AND MOTIVATION

1.1.1 The Concept of Infrastructure

Infrastructure and "social overhead capital" seems to be synonyms in the works of early development economists. For instance, as broadly defined by Rosenstein-Rodan (1943), "social overhead capital" includes roads, railways, canals, communications, and hydro-electric power stations. These "social overhead capital" are considered an essential prerequisite for industrialisation. Likewise, Nurkse (1953) suggests a massive investment in new technology, new machines and new production processes to stimulate the industrialisation process.

Hirschman (1958) goes further with a narrower definition by distinguishing between "social overhead capital" and "directly productive activities". In particular, "social overhead capital" is defined as "basic services without which primary, secondary, and tertiary productive activities cannot function" (Hirschman, 1958, p.83). Accordingly, "social overhead capital" includes a broad range of public services, such as transport, communications, power and water supply, irrigation and drainage systems, education, public health, and law and order.

In another approach, Hansen (1965) proposes two categories of infrastructure, namely, "social overhead capital" and "economic overhead capital". In this approach, infrastructures are distinguished according to the way that they affect the process of economic development in a direct or an indirect activity. As noted by Hansen (1965), "economic overhead capital" includes roads and other types of transport infrastructure, water and energy supply systems, and irrigation plants; and, "social overhead capital" includes council flats, schools, hospitals, sport structures, green areas, and structures for public safety.

According to the World Bank's the World Development Report 1994 (p.2), infrastructure is defined as "an umbrella term for many activities referred to as "social overhead capital" by such development economists as Paul Rosenstein-Rodan, Ragnar Nurkse, and Albert Hirschman. Neither term is precisely defined, but both encompass activities that share technical features (such as economies of scale) and economic features (such as spillovers from users to non-users). Economic infrastructure usually has the following characteristics, economies of scale and spillovers from users to non-users." And, the World Bank's definition of economic infrastructure is restricted to three categories, namely, public utilities, public works, and other sectors of transportation (World Bank, 1994).

Infrastructure is also categorised into "hard infrastructure" and "soft infrastructure". Bhattacharyay (2009, p.2), for example, defines the former as "physical structures or facilities that support the economy and society" and the latter as "non-tangibles supporting the development and operation of hard infrastructure". Transport, telecommunications, energy and water supply, schools, hospitals are classified as "hard infrastructure". Governance, regulatory and institutional framework, and social network are considered "soft infrastructure". According to Button (2002), there has been a shift from the definition of infrastructure which focuses only on physical infrastructure assets to the one extended to soft types of infrastructure.

Generally, infrastructure has been defined according to a wide range of approaches. In many contexts of analysing the role of infrastructure, the empirical literature has employed a variety of different indicators as proxies of infrastructure. This, in turn, poses a challenge to any comparison involving studies using different infrastructure definitions and proxies.

1.1.2 Theories and Empirical Debates on Economic Effects of Infrastructure

The World Bank (1994) refers to infrastructure as the engine of economic activity and claims that public infrastructure investment is a crucial factor driving economic growth and development. In practice, infrastructure has been a focus of government expenditures around the world. For instance, in low and middle-income countries, infrastructure typically accounts for 7 to 9 per cent of GDP and represents about 40 to 60 per cent of public investment (World Bank, 1994). Among various types of infrastructure, transport is of particular importance for development. Adam Smith expressed this as "no roads, no transport, no trade, no specialisation, no economies of scale, no productivity progress, and no development" (Prud'homme, 2004, p.9).

Research on economic effects of infrastructure is part of a broader field of academic inquiry on the role of public policies in economic development. The theoretical and empirical literature of economics has presented considerable attempts devoted to estimating the role of infrastructure in several contexts. With reference to the empirical literature, the majority of studies have been concerned with the role of infrastructure as an important driver of output and productivity growth. Another strand of empirical research seeks to explain the impact of infrastructure on the labour market, especially job creation. One other strand examines the empirical impact of infrastructure, particularly transport infrastructure, on industrial location choices. However, empirical studies belonging to these strands of research have yielded mixed conclusions.[1]

The relationship between infrastructure and economic growth has been the subject of continuing debate. Theoretically, Barro (1990) is the seminal paper in this field. To

[1] The role of infrastructure has also been discussed in other contexts, such as private investment and trade. However, as stated above, the thesis focuses on the impacts of infrastructure on growth, employment and industrial location choices.

demonstrate a potentially positive growth impact of public services, Barro (1990) develops a growth model in which government expenditure is considered as an additional input to private production. The Barro (1990) model has been extended in the ensuing literature to consider factors that might affect the growth impact of public services in either positive or negative ways. One recent example is the two-country model developed by Hashimzade and Myles (2010) in which the production process of a given country benefits from infrastructure spillovers originated from the neighbouring country. With respect to the empirical literature, the majority of studies have been based on the hypothesis that the growth contribution of infrastructure can be seen in both direct and indirect ways. Specifically, infrastructure is assumed to be an intermediate input directly introduced into the production process in the same way as labour and physical capital. In the indirect way, infrastructure performs as an augmenting factor that raises the productivity of other production inputs. Empirical evidence, however, varies dramatically, ranging from no growth impact to a statistically significant and positive impact of infrastructure on economic growth. Sources of this empirical conflict could include variations across studies in regards to empirical model specification, econometric techniques, infrastructure measurement, time and geographical dimension of data, and, especially for analyses at the sub-national level, attempts to control for the existence of either positive or negative spillovers of infrastructure across sub-national regions.

The employment impact of infrastructure, in particular transport infrastructure, is another popular strand of research. As the basic principle of competitive labour market theory states, the equilibrium levels of employment and wages are determined by the interaction between labour demand and labour supply. According to Eberts and Stone (1992), public investment on infrastructure, such as roads and highways, could alter the attractiveness of geographical areas, leading to an adjustment in the local labour demand and supply and, consequently, a change in the equilibrium levels of employment and wages in those areas. For instance, on the demand side, infrastructure improvement, particularly transport infrastructure improvement, could contribute to enhancing productivity and hence encourage expansion of existing firms as well as attract new establishments and re-allocations from elsewhere, thereby adjusting the local demand for labour; on the supply side, infrastructure improvement could enhance job accessibility as well as attract the in-migration of households that consider the availability of a reliable infrastructure system as a residential amenity, thereby adjusting the local supply of labour. From a spatial perspective, the creation of jobs in particular areas could be at the expense of employment in other areas where some of those jobs were previously located. Then, the overall employment impact of infrastructure could be ambiguous at the national level. In the empirical literature, the relationship between transport infrastructure and employment has been analysed by using different levels of data, different approaches of empirical modelling, and a variety of econometric techniques. However, the empirical question as to whether public infrastructure investment can have a positive impact on employment is still inconclusive. With respect to the spillover effect of transport infrastructure across geographical areas, the majority of empirical studies have been concerned with the impact of spillovers on output and productivity growth while only a few studies examine this issue in the context of an employment analysis.

Apart from other local attributes, transport infrastructure and, especially, market access, which can be assumed to be enhanced by transport infrastructure improvement, have been found to have a role to play in shaping the location choices of firms. Also, the empirical literature emphasises that the impact of local attributes on the location choices varies according to firm-specific characteristics, such as firm size, industrial activity, or industrial organisation. Theoretically, the relationship between transport infrastructure and firm location choices finds its root in classical location theories which emphasise the importance of transport costs. Over time, as noted by Preston (2001), the emphasis has been less on transport costs in the classical sense, while increased on transport infrastructure improvement as a factor that interacts with agglomeration economies. Particularly since the

1990s, the new economic geography has provided a further insight into industrial agglomeration which has different patterns across non-homogenous geographical areas. In this strand of research, transport infrastructure remains an important factor determining the locational behaviour of firms. However, it is necessary to recall the conclusion of Blair and Premus (1987) on various implications of different location factors at different stages of economic development. That is, because the economy needs to shift to advanced technologies to maintain its competitiveness, the importance of "non-traditional" location factors, such as business climate or labour skills, would increase while "traditional" location factors, for example labour, transport costs and distance to raw materials, might decline in their influences on the choices of firms to locate in particular areas. Similarly, although supporting the notion that firms locate their factories according to the variations in their preferences to factors that are often heterogeneous and potentially immobile across geographical areas, such as production inputs and infrastructure, Hayter (1997) concludes that a much greater emphasis has been given to such factors as public policies, business strategies and structures, the structure of labour markets, as well as to the relationships between these factors, to examine the dynamics of industrial location. The arguments of Blair and Premus (1987) and Hayter (1997) mentioned above lead to a question as to whether transport infrastructure remains an important factor explaining why firms choose to locate in particular areas, but not elsewhere, given that transport infrastructure improvement is assumed to be positively associated with an improvement in market access.

The briefly discussed theoretical and empirical background of the economic effects of infrastructure suggests that how infrastructure causes growth, how infrastructure stimulates employment and how infrastructure influences the locational behaviour of firms remain open issues for further research. An analysis with a reference to the Vietnamese economy can be viewed as an empirical contribution to this line of research. On the one hand, the literature related to the Vietnamese economy is still limited, and on the other hand, Vietnam can be a good illustration for the argument that infrastructure investment is one approach to development.

1.1.3 The Vietnamese Economy during the Period 2000-2007[2]

The 1986 Renovation represents a significant change in the way the State of Vietnam approaches economic issues. Its main contents are cited in GSO (2004, p.1033), which can be summarised as follows: (i) application of the mechanism of a market economy with the governing role of the State, where the State would govern the operation of this market economy by a legal system; (ii) transformation from an economy in which state-owned enterprises and collectives had a dominant role to an economy with a multi-sectoral structure of economic agents, and transformation from the economic mechanism operated by production-input allocation and state subsidises to the mechanism of value-based production with autonomy in business activities; and (iii) diversification of external economic relations. Subsequent economic reforms have been based on these key contents of the 1986 Renovation.

The most direct indicator of economic progress would be economic growth. During the period 2000-2007, which is the sample period of the empirical analyses of the thesis, the economy achieved an average real Gross Domestic Product (GDP) growth rate of 7.63 per cent per year. Growth can also be observed in all economic sectors during this period. Specifically, industry grew at an average rate of 10.24 per cent per year, playing the role as the leading sector of the economy. The average growth rate of the service sector was 7.16 per cent per year, relatively lower than that of the industry sector. The average growth rate of

[2] This section describes briefly the economy during the period 2000-2007 which is the period examined in the empirical analyses of the thesis. A general background on the Vietnamese economy is presented in Chapter 2. The focus of the empirical analyses on the period 2000-2007 is essentially driven by the good availability of province-level data for conducting econometric experiments. The sources of data cited in this section are the General Statistics Office of Vietnam (GSO) and Asia Development Bank (ADB), which are to be clarified in Chapter 2.

3.90 per cent per year of the agriculture sector was the lowest growth rate among the three sectors.

Although the annual proportional change of the agriculture sector was positive during the period observed, its share in total GDP declined from 23.28 per cent in 2000 to 17.93 per cent in 2007. Over this period, the non-agricultural sectors accounted for an average share of 79.35 per cent per year in total GDP. These figures illustrate the on-going process of structural transformation towards industrialisation of the economy. In relation to the role of foreign-invested enterprises in the Vietnamese economy, the contribution of the foreign-invested sector to total GDP increased from 6.73 per cent in 1995 to 7.16 per cent in 2000 and then 13.29 per cent in 2007.

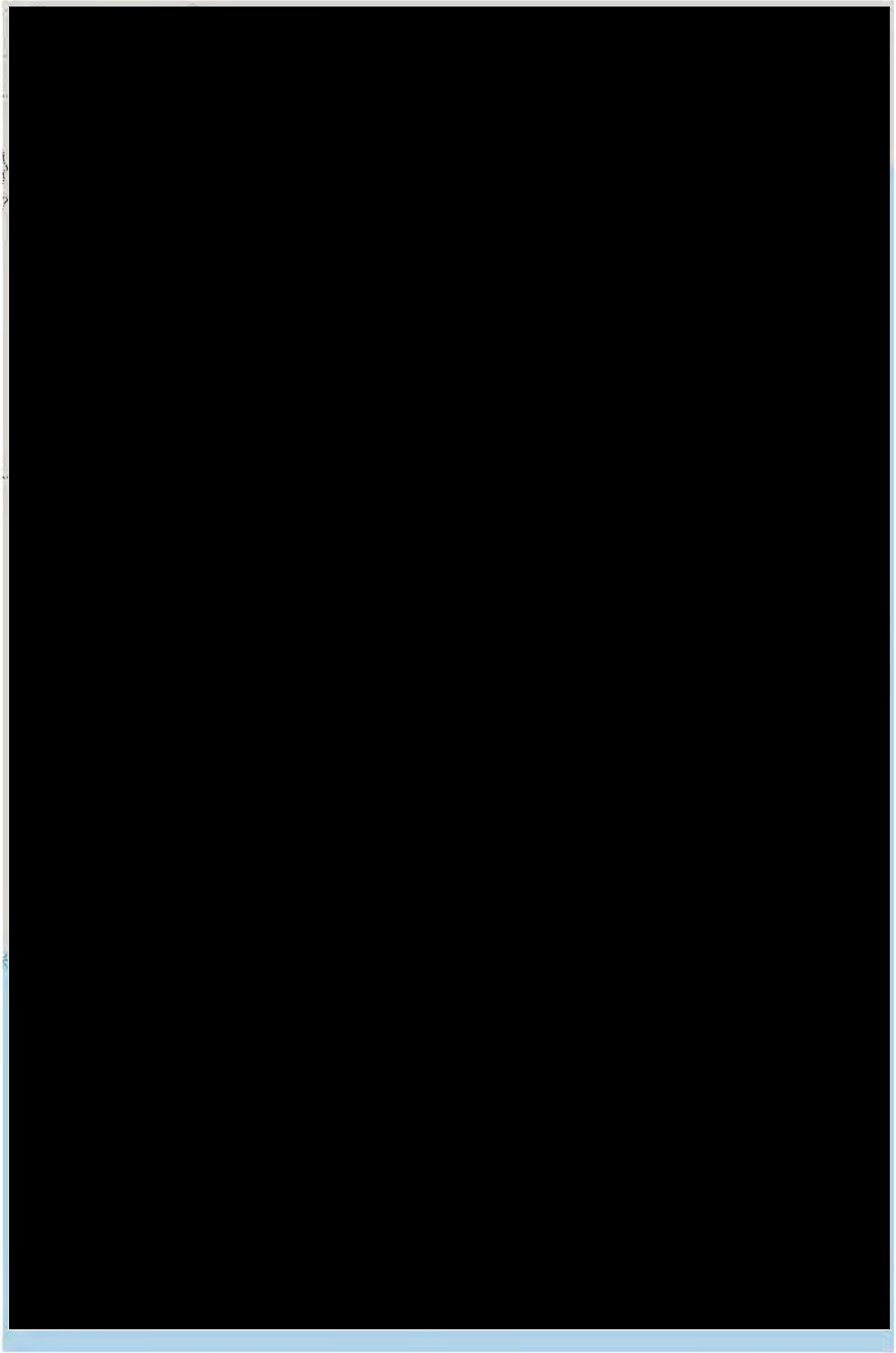
The ratio of exports plus imports to GDP increased from 54.13 per cent in 1990 to 91.49 per cent in 2000 and then 151.3 per cent in 2007, illustrating Vietnam's progress in trade openness. There was a transformation in regard to the structure of exports towards an increase in the proportion of non-agricultural products during this period. Specifically, the share of agriculture in total exports declined from 46.27 per cent in 1995 to 28.98 per cent in 2000 and then 23.07 per cent in 2007. The share of foreign-invested enterprises in total exports increased from 27.04 per cent in 1995 to 47.02 per cent in 2000 and then 57.19 per cent in 2007.

The period 2000-2007 also witnessed the explosion of private entrepreneurship. The most striking evidence would be the development of private sector in terms of new establishments. The state sector accounted for the average share of 6.62 per cent per year in total number of enterprises, which was much smaller than the average share of 89.89 per cent per year of the non-state sector but about twice the average share of 3.49 per cent per year of the foreign-invested sector. While the non-state sector and the foreign-invested sector achieved the high average growth rates of 22.88 per cent per year and 18.49 per cent per year, respectively, the state sector experienced a decline at the average rate of -6.83 per cent per year, in its number of enterprises. The foreign-invested sector was dominated by 100 per cent foreign-owned enterprises which accounted for the average share of 71.26 per cent per year of total number of foreign-invested enterprises and grew at the average rate of 25.20 per cent per year. In regard to job creation, the state sector accounted for the average share of 41.12 per cent per year of total number of employees; but, its workforce was declining annually at the average rate of -2.24 per cent. Non-state enterprises and foreign-invested enterprises experienced the high average growth rates of 21.02 per cent per year and 22.73 per cent per year, respectively, in their number of employees, thus playing an increasingly important role in creating jobs.

Infrastructure has been an important component of Vietnam's development strategy and, therefore, the focus of public spending. The World Bank (2000, p.1) notes that "the dramatic increase in the provision of infrastructure services in Vietnam since the late 1980s greatly facilitated rapid growth in exports and GDP." During the period 2000-2007, a large proportion of public spending was allocated to infrastructure investment. Specifically, the ratio of public infrastructure investment to GDP in was close to 10 per cent in 2007, which is viewed as very high by international standards (World Bank, 2009).

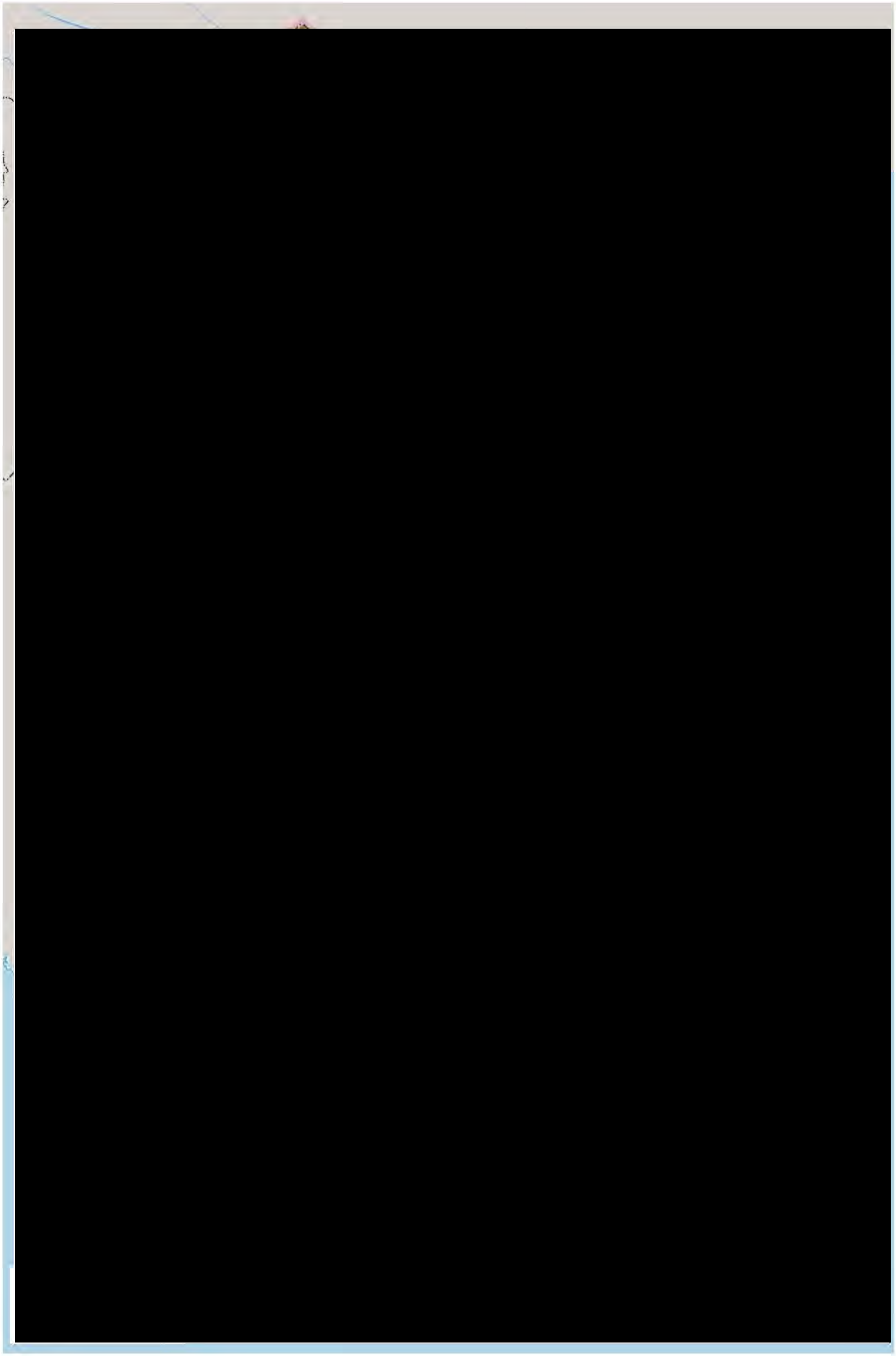
From the spatial perceptive, the economy of Vietnam has been characterised by an uneven distribution of resources and economic progress across provinces and regions. According to various publications of the General Statistics Office of Vietnam, provinces are grouped into six regions, namely, the Red River Delta, the Northern Mountain, the Central Coast, the Central Highlands, the Southeast and the Mekong River Delta. For the purposes of demonstrating Vietnam's geography from provincial and regional dimension, Figures 1.1 and 1.2 present the map of Vietnam and the map of Vietnam by regions.

Figure 1.1: Map of Vietnam



Source: Based on the Vietnamese-language version published by the Cartographic Publishing House of Vietnam (2004).

Figure 1.2: Map of Vietnam by Regions



Source: Based on the Vietnamese-language version published by the Cartographic Publishing House of Vietnam (2004).

The level of economic development has been very different between regions. One indicator of the economic development level can be the average per capita income of the region relative to the national average.[3] Over the period observed, the Southeast had the highest average per capita income which was more than twice the national average. This is followed by the Red River Delta and the Mekong River Delta where per capita incomes were approximately equal to the national average; whereas, the average per capita incomes in other regions were lower than the national average. The poorest was the Northern Mountain which had the lowest average per capita income among the six regions. Geographical disparity in economic development remains an issue to be solved, and the economic literature says that investment in infrastructure is a useful tool for narrowing the development gap geographically. However, the level of infrastructure development has varied dramatically throughout the country. For instance, during the period 2000-2007, the average road density, measured as the average road length per 1,000 km^2, was highest in the Red River Delta, followed by the Southeast. In regard to education infrastructure, the majority of higher-education institutions were located in these two regions as well. Generally, the endowment of infrastructure in other regions lagged behind that in the Red River Delta and the Southeast. At the province level, Ha Noi and Ho Chi Minh City have been the largest economic centres of Vietnam. During the period 2000-2007, while the former accounted for an average share of 7 per cent per year in the national real GDP, the latter was the largest contributor to the national real GDP with an average contribution of 18 per cent per year.

Since 1986, reforms have led to a number of important economic achievements in Vietnam. Within the context of the thesis, several important achievements have been briefly described above. In particular, during the period 2000-2007, the economy achieved a high growth rate of GDP, a fundamental transformation in its GDP structure towards industrialisation, an

[3] The thesis considers the per capital GDP level an indicator of the economic development level. This is based on the idea of Lucas (1988, p.3), stating that "By the problem of economic development I mean simply the problem of accounting for the observed pattern, across countries and across time, in levels and rates of growth of per capita income. This may seem too narrow a definition, and perhaps it is, but thinking about income patterns will necessarily involve us in thinking about many other aspects of societies too …".

impressive growth of external trade, and an explosion of private entrepreneurship. The development of the private sector was most impressive in terms of the increased number of non-SOEs, i.e. non-state enterprises and foreign-invested enterprises, and hence the increased proportion of jobs created by those enterprises. These development outcomes had been observed at the same time as a strong focus of public spending on infrastructure.

1.2 TERMINOLOGIES, RESEARCH QUESTIONS AND METHODOLOGY

The thesis is generally aimed at providing an empirical contribution to understanding of the economic role of infrastructure in Vietnam. In particular, the empirical research examines three broad aspects, namely, the impact of infrastructure on economic growth, on private employment and on the locational behaviour of new enterprises belonging to the private sector. It is, first of all, important to clarify several terms used in the thesis.

1.2.1 Terminologies

The three empirical studies of the thesis focus on two categories of infrastructure: roads, measured as road length per 1,000 km^2 of the provincial area, and higher-education institutions, proxied by the number of colleges and universities. The former is also referred to as transport infrastructure throughout the empirical analyses. The latter represents higher-education infrastructure. Infrastructure enters into the empirical models as a stock variable. Due to the unavailability of province-level information on charges/ fees of, and congestion in, using infrastructure, infrastructure is assumed to be non-excludable and non-rivalrous.

In various publications of the General Statistics Office of Vietnam, from a perspective of ownership, the economy of Vietnam contains three broad sectors: the state sector, the non-state sector and the foreign-invested sector. In the thesis, for the purposes of analytical and empirical analyses, the term the private sector is used to refer to the sector containing both the non-state sector and the foreign-invested sector.

The thesis also uses data from the Annual Enterprise Surveys which have been conducted by the General Statistics Office of Vietnam since 2000. There are a number of terminological issues with respect to this database which should be made clear. Firstly, “enterprise” is defined as "an economic unit that independently keeps a business account and acquires its own legal status" (GSO, 2010, p.19). Accordingly, the Surveys exclude registered enterprises which have not begun operations, enterprises which have been disbanded, and economic units which do not keep business accounts, such as branches. Secondly, a distinction is maintained between different categories of enterprises. The state sector comprises enterprises with 100 per cent of state capital, public limited liability companies, and companies with the State's shares greater than 50 per cent in total registered capital. The non-state sector comprises cooperatives, private companies, private limited liability companies, private stock companies and stock companies with a less than 50 per cent of the State's share in total registered capital. The foreign-invested sector comprises joint ventures and enterprises with 100 per cent capital owned by foreigners.[4]

Administratively, at the sub-national level, Vietnam is divided into provinces. From 2000 to 2003, the total number of provinces was 61. In 2004, there were 64 provinces as a result of the division of 3 provinces into 6 new provinces. After 2007, there have been 63 provinces. For consistency in the total number of cross-sections of the panel dataset, and to ensure that all provinces are examined in the empirical research, the provinces divided in 2004 were re-joined in the sense that data of each pair of divided provinces were summed up to generate data of the origin province. Therefore, the total number of provinces examined in the thesis is 61.

For the purposes of analysing analytically the geographical distribution of resources and economic activities, the thesis also categorises provinces into different regions according to their geographical location. This is based on the regional classification of the General Statistics Office of Vietnam. As mentioned previously, there are six regions, namely, the Red

[4] Enterprise shall be established and shall operate under the regulation of the State Enterprise Law, Cooperative Law, Enterprise Law, or Foreign Investment Law (GSO, 2010). See these legal documents for a definition of each type of enterprise.

River Delta, the Northern Mountain, the Central Coast, the Central Highlands, the Southeast and the Mekong River Delta. This classification is essentially geography-based, capturing regional differences in terms of climate, terrain and natural resources.

With respect to the term metropolis, there are 6 types of metropolis: special metropolis; type-I, type-II, type-III, type-IV, and type-V metropolises.[5] As of 2007, the two special metropolises and the three type-I metropolises are distributed vertically from the North to the South. Ha Noi, the special metropolis, and Hai Phong, the type-I metropolis, are located in the northern region, i.e. the Red River Delta. Da Nang, the type-I metropolis, is located in the central region, i.e. the Central Coast. Ho Chi Minh City, the special metropolis, is located in the southern region, i.e. the Southeast. Finally, Can Tho, the type-I metropolis, is located in another southern region, i.e. the Mekong River Delta.

1.2.2 Research Questions

To examine the economic role of infrastructure, the thesis focuses on the above-stated issues, namely, economic growth, private sector employment and private sector location choices. Research questions can be grouped into three broad themes as follows.

- The impact of infrastructure on the growth of per capita GDP, on the employment in the private sector, and on the locational behaviour of new enterprises belonging to the private sector;

[5] As defined in Decree No. 72/2001/NĐ-CP of the Government of Vietnam, a special metropolis is (i) the Capital of Vietnam or an area that is the centre for political, economic, cultural, science-technology, education, tourism, services activities, the national and international transport hub and the engine for socio-economic development of the whole country; (ii) has more than 90 per cent of labour working in non-agro industries; (iii) has a fully-functional and basically developed infrastructure network; (iv) has a population of more than 1.5 million; and (v) has an average population density of more than 15 thousand persons/ km2. A type-I metropolis is (i) an area that is the centre for political, economic, cultural, science-technology, tourism, services activities, the national and international transport hub and the engine for socio-economic development of the region where she is located or the whole country; (ii) has more than 85 per cent of labour working in non-agro industries; (iii) has a basically developed infrastructure network; (iv) has a population of more than 500 thousands; and (v) has an average population density of more than 12 thousand persons/ km2. See Decree No. 72/2001/NĐ-CP for the definitions of other types of metropolis.

- The existence of spillovers from transport infrastructure across neighbouring provinces;
- The sectoral variations in regard to the impact of infrastructure.

1.2.3 Methodology

Panel data analyses are undertaken in the three empirical chapters. Econometric regression approaches are employed to investigating the above-stated research questions. The empirical analyses involve estimating econometric models specified as to isolate statistically the impact of infrastructure from those caused by other factors suggested by the relevant theoretical and empirical literature. The specification of the empirical models is in accordance with the relevant literature and the context of the Vietnamese economy, and subject to the nature and the availability of data.

1.3 OUTLINE OF THE THESIS

The thesis contains 6 chapters, including the present one. The outline of the remaining chapters is presented below.

Chapter 2 provides a background discussion on the pre-1986 and post-1986 periods of the Vietnamese economy. The chapter begins with a description of the economy during the pre-reform period, i.e. from 1976 to 1985, focusing on factors driving the economy into the serious crisis at the end of this period. The remainder of the chapter is all about the post-1986 economy. First of all, the 1986 Renovation and subsequent economic reforms are presented in order to establish a brief description of the legal framework in which the whole economy, and especially the private sector, has since operated. This is followed by a discussion on several aspects of the post-1986 economy, namely, GDP growth, macroeconomic stability, structural transformation of GDP, growth and diversity of external trade, evolution of the labour market, and explosion of private entrepreneurship. Subsequently, an illustration of physical infrastructure development is presented with an emphasis on transport infrastructure and education infrastructure. Next, the spatial economy of Vietnam is described through a discussion on the geographical distribution of resources and the provincial differences in terms of economic structure, export activity, private sector development and infrastructure provision over the period 2000-2007. The final section summarises key aspects that together establish a general background on the economy observed in the subsequent empirical chapters.

Chapter 3 is the first empirical chapter of the thesis, examining the impact of infrastructure on economic growth. The literature review is concerned with the current state of knowledge regarding the relationship between infrastructure and economic growth. Theoretically, the review focuses on the Barro (1990) model and its ensuing extensions. The review of the empirical literature emphasises econometric issues associated with estimation of the panel-data growth regression, shortcomings of monetary and physical infrastructure measuring approaches, implications of specific characteristics across regions and sectors for the growth impact of infrastructure, and spillover effect of transport infrastructure across sub-national areas. To some extent, these empirical aspects could provide some explanations for the contradicting results obtained in previous studies. The subsequent section presents an empirical growth model for the Vietnamese provinces, concentrating on the growth impact of infrastructure. Empirical results are obtained by using the System Generalised Method of Moment (System-GMM) estimator which appears to be the most viable option for the growth analysis with the presence of the lagged level of the dependent variable and with the inclusion of other explanatory variables, such as infrastructure, that can give rise to the problem of endogeneity. Furthermore, as Blundell and Bond (1999) and Bond *et al.* (2001) suggest, the System-GMM estimator is more efficient than the Difference-GMM estimator for a panel with short time dimension and persistent series. In addition to identifying infrastructure as an engine of growth, the empirical analysis examines the existence of cross-province transport infrastructure spillovers and provides an insight into how GDP growth was determined by other factors, such as the lagged level of GDP, physical capital investment, human capital supply, macroeconomic stability, structural transformation and external trade.

Chapter 4 examines the impact of infrastructure on the employment in the private sector. The literature review includes a discussion on how infrastructure can affect the labour market, and a summary of empirical findings. In the empirical analysis, this chapter develops two versions of employment model for the Vietnamese provinces. The first version examines the impact of infrastructure, among other factors, on the private sector employment. The second version is an extension of the first one, which adds a variable measuring transport infrastructure installed in neighbouring provinces to examine the existence of cross-province infrastructure spillovers. With respect to methodology, the Feasible Generalised Least

Squares (FGLS) estimator is applied with a control for heteroskedasticity and panel-specific first-order autocorrelation.

Chapter 5 is the final empirical chapter of the thesis, examining the impact of infrastructure on the location choices of new enterprises belonging to the private sector. Theoretically, the literature review is dedicated to the classical location-choice models and the new economic geography. It then provides a survey of empirical studies relevant to the relationship between transport infrastructure and industrial location choices. The empirical analysis presented in this chapter adopts the count data modelling approach, which is preferable over the discrete choice modelling approach since being viewed as helpful to understanding why some particular provinces are chosen to be the locations of enterprises while the others were not chosen by any enterprises (Guimarães *et al.*, 2004). Due to the nature of data, the Negative Binomial Regression technique is applied to estimate the count data regression model. The empirical analysis is then extended to examine the existence of cross-province spillover effect of transport infrastructure. Disaggregate analyses are also performed to obtain information that cannot be shown in the aggregate analysis. Specifically, the empirical location-choice models are estimated for not only the full-sample, but also sub-samples in which enterprises are distinguished according to their workforce size and industry. This is aimed at testing whether the location preferences are also dependent on specific characteristics of enterprises.

Chapter 6 concludes the thesis, containing a summary of empirical findings and their policy implications, an acknowledgement of limitations of the empirical studies and then a discussion on feasible areas for future research.

CHAPTER 2

GENERAL BACKGROUND ON THE VIETNAMESE ECONOMY

Chapter 2 begins with an overview of the economy during the pre-reform period, i.e. from 1976 to 1985, focusing on factors driving the economy into a serious crisis at the end of this period. The rest of Chapter 2 is about the post-1986 economy. Section 2.2 discusses briefly the 1986 Renovation and subsequent economic reforms, and then describes the post-1986 economy, focusing on the following main aspects: GDP growth, macroeconomic stability, structural transformation of GDP, growth and diversity of external trade, evolution of the labour market, and explosion of private entrepreneurship. Section 2.3 describes the development of physical infrastructure. The spatial distribution of resources as well as spatial differences in terms of economic performance, private sector development, infrastructure provision, and so forth, are presented in Section 2.4. Finally, Section 2.5 concludes this chapter. It is necessary to make clear that this chapter is not aimed at providing a comprehensive discussion on the Vietnamese economy. Instead, it presents a general background on the economy observed in the empirical analyses of the thesis.

2.1 THE PRE-1986 ECONOMY

The following description of the pre-1986 economy is essentially based on the values calculated as average per year for the period 1976-1985. As shown in Table 2.1, the average annual growth rate of gross social products (GSP) was 3.78 per cent.[6] However, in some years, the national economy experienced negative growth rates.

In regard to ownership structure, expressed by the share of each category of economic agents in GSP, the household sector dominated the economy with an average GSP share of 45.32 per cent per year, followed by the collective sector. The average GSP share of the State-Owned Enterprise (SOE) sector was smaller than those of households and collectives. Although holding the largest average GSP share, the household sector experienced an average negative growth rate during this period. With respect to economic growth, the highest GSP growth rate can be seen in the collective sector. While the SOE sector experienced an average positive growth rate, its growth rate was much lower than that of the collective sector. Also, the collective sector always maintained a positive GSP growth rate over the period observed. Comparing the performance between the three sectors, particularly from 1980 to 1985, the collective sector was likely to be the most important contributor to the positive GSP growth rate of the national economy, although it did not hold the largest share of GSP.

Given that productivity can be proxied by GSP per employee, presented in Table 2.2, the collective sector was the most inefficient sector, compared with the SOE and household sectors. However, the largest proportion of jobs was created by collectives. There was no evidence of significant changes of population growth rate, unemployment rate and the labour share of SOEs; in contrast, a significant decrease in the labour share of households can be observed from 1981 to 1985 while, during the same period, collectives experienced a

[6] The term gross domestic product (GDP) did not appear as a statistical indicator for the pre-1986 period in GSO (2004).

significant increase in their labour share. As presented in GSO (2004), to fulfil the 1976-1980 five-year plan which set quantity-based targets for agricultural and industrial production over this period, several measures had been taken to promote the expansion of the SOE and collective sectors.

With respect to the sectoral structure, agriculture was the most important sector of the economy, contributing the largest proportion of an average 50.44 per cent per year to GSP. However, its average growth rate was only 4.06 per cent per year and the growth rate fluctuated between 1976 and 1985. Industry achieved a relatively higher average growth rate compared with that of agriculture, i.e. an average growth rate of 4.98 per cent per year. Nevertheless, its GSP share was approximately one half of that of agriculture. Concerning the productivity of these sectors, industry had the lowest level of GSP per employee. Despite the fact that there were some changes in the allocation of state investments towards the industry and construction sectors, the agriculture-dominated economic structure was relatively persistent over this 10-year period. This could indicate that resources were not allocated and used in an efficient way and, consequently, further investments were unable to raise the role of non-agricultural sectors in the economy.

Table 2.1: GSP Structure, GSP Growth and Inflation (1976-1985)

Year	GSP Growth (%)	GSP by Economic Agents (%)						GSP by Economic Sectors (%)								Retail Price Index (%)
		SOEs		Collectives		Households		Agriculture (incl. Forestry)		Industry		Construction		Others		
		GSP Share	GSP Growth	GSP Share	GSP Growth	GSP Share	GSP Growth	GSP Share	GSP Growth	GSP Share	GSP Growth	GSP Share	GSP Growth	GSP Share	GSP Growth	
1976	...	25.00	...	24.28	...	50.72	...	50.22	...	26.05	...	3.98	...	19.75	...	121.9
1977	2.82	26.01	6.99	25.17	6.61	48.82	-1.05	48.32	-1.07	28.30	11.73	3.95	1.85	19.43	1.17	118.6
1978	2.27	28.11	10.52	25.19	2.35	46.70	-2.16	45.70	-3.27	30.27	9.37	3.81	-1.31	20.23	6.44	120.9
1979	-2.02	26.43	-7.86	26.29	2.26	47.27	-0.82	47.38	1.57	29.24	-5.35	3.81	-1.94	19.58	-5.17	119.4
1980	-1.39	23.03	-14.11	29.65	11.20	47.32	-1.28	51.05	6.26	26.03	-12.21	3.69	-4.57	19.23	-3.14	125.2
1981	2.33	22.85	1.55	30.18	4.15	46.97	1.57	52.22	4.68	25.58	0.54	3.44	-4.51	18.76	-0.17	169.6
1982	8.83	21.48	2.31	31.93	15.15	46.59	7.95	53.06	10.58	25.77	9.67	2.68	-15.13	18.48	7.24	195.4
1983	7.15	20.45	2.02	33.81	13.46	45.74	5.20	53.54	8.13	25.68	6.76	2.85	13.93	17.93	3.92	149.4
1984	8.31	22.43	18.79	37.17	19.06	40.40	-4.32	51.53	4.25	26.84	13.20	2.97	12.71	18.66	12.76	164.9
1985	5.69	24.37	14.84	42.91	22.02	32.72	-14.42	51.42	5.45	28.22	11.14	3.06	8.89	17.31	-2.00	191.6
Average	3.78	24.02	3.89	30.66	10.70	45.32	-1.04	50.44	4.06	27.20	4.98	3.42	1.10	18.94	2.34	...

Source: Retail Price Index is taken directly from GSO (2004); whereas, others are calculated by using data from GSO (2004).
Note: Gross Social Products (GSP) is expressed in 1982 constant prices.

Table 2.2: Population, Employment and Labour Productivity (1976-1985)

Year	Popula-tion Growth (%)	Labour and GSP p.e. by Economic Agents (%)						Labour and GSP p.e. by Economic Sectors (%)								Unemp-loy-ment (%)
		SOEs		Collectives		Households		Agriculture (incl. Forestry)		Industry		Construction		Others		
		Labour Share	GSP p.e.	Labour Share	GSP p.e.	Labour Share	GSP p.e.	Labour Share	GSP p.e.	Labour Share	GSP p.e.	Labour Share	GSP p.e.	Labour Share	GSP p.e.	
1976	...	...	...	...	...	...	...	66.86	14.98	11.20	4.84	4.80	5.35	17.14	7.42	0.21
1977	2.55	...	...	...	...	...	...	67.87	16.32	11.06	4.54	5.00	5.04	16.07	7.71	0.21
1978	2.00	...	...	...	...	...	...	68.61	17.16	11.02	4.16	5.27	4.52	15.10	8.37	0.21
1979	2.02	...	...	...	...	...	...	69.23	15.43	11.11	4.01	5.09	4.39	14.58	7.87	0.20
1980	2.40	...	...	...	...	...	...	70.71	13.90	10.40	4.01	4.66	4.39	14.23	7.50	0.18
1981	2.24	14.94	8.35	54.67	3.01	30.39	8.43	70.62	12.96	10.77	4.04	4.14	4.53	14.47	7.08	0.19
1982	2.26	14.25	8.57	55.81	3.25	29.94	8.84	70.92	13.74	10.66	4.25	3.70	4.12	14.72	7.13	0.18
1983	2.14	13.90	8.66	55.95	3.56	30.15	8.93	71.50	13.52	11.18	4.41	3.38	4.97	13.95	7.56	0.19
1984	2.23	14.85	9.34	63.09	3.64	22.06	11.33	72.31	15.52	10.69	4.41	3.17	5.79	13.83	8.34	0.19
1985	2.08	14.72	10.44	71.53	3.78	13.75	15.01	72.94	16.54	10.76	4.45	3.20	6.04	13.11	8.33	0.19
Average	2.21	14.53	9.07	60.21	3.45	25.26	10.51	70.16	15.01	10.89	4.31	4.24	4.91	14.72	7.73	0.20

Source: Unemployment is taken directly from GSO (2004); whereas, others are calculated by using data from GSO (2004).
Note: Gross Social Products (GSP) is expressed in 1982 constant prices.

Table 2.3: Growth and Composition of State Investment (1976-1985)

Year	Growth (%)	Investment by Economic Sectors (%)							
		Agriculture (incl. Forestry)		Industry		Construction		Others	
		Share	Growth	Share	Growth	Share	Growth	Share	Growth
1976	...	22.29	...	31.91	...	5.36	...	40.44	...
1977	24.88	26.89	50.61	31.34	22.63	5.05	17.61	36.73	13.42
1978	9.27	26.09	6.02	33.65	17.34	5.62	21.66	34.64	3.06
1979	-2.46	22.98	-14.06	39.14	13.47	4.41	-23.50	33.46	-5.78
1980	-6.38	21.79	-11.26	40.72	-2.62	5.43	15.16	32.07	-10.27
1981	-10.08	22.65	-6.50	46.29	2.23	3.47	-42.49	27.59	-22.65
1982	0.40	18.19	-19.38	47.39	2.80	2.65	-23.29	31.76	15.61
1983	28.66	19.13	35.33	40.38	9.62	3.55	72.51	36.93	49.60
1984	25.28	22.05	44.38	33.76	4.74	2.29	-19.12	41.90	42.12
1985	6.94	21.72	5.33	31.21	-1.14	2.07	-3.56	45.01	14.88
Average	8.50	22.38	10.05	37.58	7.67	3.99	1.66	36.05	11.11

Source: Calculation using Data from GSO (2004)
Note: Investment is expressed in the 1982 constant prices.

Increases in the inflation rate had become uncontrollable since 1976, as GSO (2004) notes. By 1985, the failure of the "Price-Salary-Money" reform, which introduced "a shock remedy" into the economy, caused the inflation rate to "skyrocket" (GSO, 2004, p.37). As described by GSO (2004, p.37), "prices escalated daily … not only in the free market but also in the organised one," and "basically, prices got completely out of the control of the State." This period also experienced both an increased budget deficit which was 18.1 per cent and 36.6 per cent in 1980 and 1985 respectively and had to be financed by printing money (GSO, 2004). Also, as can be seen in Table 2.4, there had been an increasing trade deficit over the same period. After all, the statistical indicators of inflation, public finance and trade balance could illustrate the macro instability of the economy.

As can be seen in GSO (2004), during the period 1976-1985 Vietnam's trading partners were statistically categorised into two main groups: socialist countries and others, including non-socialist countries and international organisations. Trade deficit with socialist countries had expanded annually. In particular, the value of trade deficit with socialist countries increased approximately 3-fold from 1976 to 1985. Both exports and imports between Vietnam and socialist countries experienced positive growth rates in most years during this period. However, as presented in GSO (2004), the value of exports was much lower than that of imports and, thereby, it could cover only about 30 per cent of imports. It is worth noting that almost all essential materials for production and basic commodities for household consumption such as rice and fabrics, which could had been produced domestically, were imported completely or partially, and this was essentially because domestic production was unable to satisfy demand (GSO, 2004). All indicators of trade relations with other partners fluctuated dramatically from year to year, illustrating the unstable economic relationship between Vietnam and non-socialist countries during this pre-reformed period.

Table 2.4: Exports, Imports and Trade Balance (1976-1985)

Year	Trade with Socialist Countries			Trade with Other Trading Partners		
	Export Growth (%)	Import Growth (%)	Trade Balance (thousand Rubles)	Export Growth (%)	Import Growth (%)	Trade Balance (thousand USD)
1976	...	...	-341446	...	...	-356810
1977	77.91	18.38	-336272	12.92	21.64	-441858
1978	11.65	-7.76	-270765	-21.01	21.33	-578990
1979	3.66	60.71	-564487	6.90	-15.26	-472882
1980	-3.75	-5.55	-528951	-11.17	-13.49	-407092
1981	4.17	25.52	-712222	100.67	-10.08	-281909
1982	43.09	14.78	-750765	20.11	-11.51	-201240
1983	13.10	4.84	-759254	28.07	0.47	-151622
1984	6.99	8.07	-824712	2.37	32.69	-272306
1985	4.38	14.23	-982274	11.58	-12.30	-181480
Average	17.91	14.80	NA	16.72	1.50	NA

Source: Calculation using data from GSO (2004)

As GSO (2004, p.27) concludes, the post-wartime economy of Vietnam "was badly hit by successive failures of the two 5-year plans," and "by the end of 1985, the economy completely fell into crisis." Hyperinflation reached its peak in 1986 when an annual price increase was up to 774.7 per cent. In regard to the reason for the mid-1980s economic crisis, it is acknowledged that "the key factor was the highly centrally planned economic mechanism … The economy was characterised by an inefficient economic structure, massive and unplanned development of the state economic sector, prioritising heavy industry and large-scale production at the expense of small-scale production, and neglecting of market laws" (GSO, 2004, p.38).

From the policy perspective, it is also worth noting that shortcomings of the centrally-planned economic system had already been recognised by 1979. As the 1979 Resolution cited in GSO (2004, p.34) states, "the economic plans were made on the bureaucratic central planning." The 1979 Resolution is considered "a first move of great importance in gradually shifting the economy from the highly centrally-planned management mechanism and paving the way for economic reforms in the succeeding years". Some further efforts had been made during the period 1981-1985 (see GSO, 2004). However, the bad performance of the economy, as illustrated by the statistical figures presented above, might suggest that the first moves of economic reforms were not sufficient, or the expected outcomes of economic transition could not occur in the short time.

2.2 THE POST-1986 ECONOMY

This section describes the performance of the economy from 1986 to 2010. It briefly presents the 1986 Renovation and subsequent reforms, and then describes the post-1986 economy, focusing on GDP growth, macroeconomic stability, structural transformation, growth and diversity of external trade, evolution of the labour market, and explosion of private entrepreneurship.

2.2.1 The 1986 Renovation and Subsequent Reforms

The 1986 Renovation represents a significant change in the way the State of Vietnam approaches economic issues. Its main contents are cited in GSO (2004, p.1033), which can be summarised as follows: (i) application of the mechanism of a market economy with the governing role of the State, where the State would govern the operation of this market economy by a legal system; (ii) transformation from an economy in which state-owned enterprises and collectives had a dominant role to an economy with a multi-sectoral structure of economic agents, and transformation from the economic mechanism based on production-input allocation and state subsidises to the mechanism of value-based production with autonomy in business activities; and (iii) diversification of external economic relations. For the first point, the State is re-defined as "an administrative apparatus rather than a supreme economic player" (Vuong, 2010, p.2). For the second point, it recognises the legitimate existence of the private sector in the economy; and later, private ownership was officially recorded in the 1992 Constitution. For the third point, the Vietnamese economy has been opened widely to foreign trading partners and investors regardless of their country of origin. Subsequent economic reforms have been based on these key contents of the 1986 Renovation. The next paragraphs summarise several important economic reforms that have been carried out since 1986, focusing on the reforms that are relevant to the research questions of the thesis.

During the late 1980s, several important reforms, as highlighted in GSO (2004), should be mentioned here. Firstly, further autonomy in agricultural production was granted to farmers, thus encouraging further expansion in workforce and further investment in physical capital in order to expand further production. Secondly, SOEs were granted autonomy in doing business, and the State would no longer cover the losses made by these SOEs.[7] Thirdly, the legal environment for the operation of private enterprises was created by a number of legal documents. Fourthly, for the first time, the Law on Foreign Investment, with numerous incentives for foreign investors, was promulgated in 1987. In addition, to promote exports, the State gradually eased its control on export quotas. Also, the domestic currency, i.e. the Vietnam dong (VND), was devalued against the U.S. dollar, and the exchange rate became relatively close to the free market's exchange rate by 1989. Furthermore, prices of commodities were let to float in the market. Finally, commercial banks were allowed to set up; foreign banks were allowed to open their branches in Vietnam; and gold, silver and jewellery business activities were liberalised to all economic agents.

During the 1990s and 2000s, there were several legal documents promulgated to create a legal environment supportive to business activity of the private sector. Some of them are mentioned here. Firstly, the 2000 Enterprise Law was an attempt to create a level playing field for all enterprises regardless of their ownership and, in particular, presented an important support to promoting the private sector development. According to Van Arcadie and Mellon (2004, p.172), "the Enterprise Law codifies mechanisms to protect the rights of

[7] According to GSO (2004), reforms of the state sector began in 1989. At the initial stage, many small SOEs were merged with larger ones, leading to a fall in the number of SOEs. Since the mid-1990s the process of equitisation has been implemented, which transfers the state ownership into shares that are then sold but also can be partially kept by the State.

citizens to establish and operate private businesses." The key effects of the Enterprise Law include simplified business registration procedures, reduced the associated time and clarified investor protection mechanisms. Secondly, the 1987 Foreign Investment Law was amended in 1992, 1996, 2000 and finally replaced by the Investment Law which went into effect in 2006. The 2006 Investment Law is the legal document which creates a unified regulatory and legal framework for all investors, regardless of their nationality, and all enterprises, regardless of their type of ownership. Thirdly, according to O'Conner (1996), the Labour Code, which went into effect in 1995 and then amended in 2002 and 2006, provides a common legal framework for employment relations in both the public and the private sector, and this should contribute to accelerating the process of labour market integration.

The above presentation of several legal reforms is aimed at supporting the subsequent discussion on the economic background for the empirical research performed in the thesis. Therefore, it is obviously not enough to illustrate the progress of legal reforms in Vietnam. However, as Menon (2009) notes, although the Vietnamese legal reforms have gained remarkable achievements in improving the business climate, further improvement is still required. This could lead to the conclusion that the legal system required for Vietnam's market economy to operate has not been well-developed.

2.2.2 GDP Growth and Macroeconomic Stability

As can be seen in Table 2.5, the economy achieved an average annual GDP growth rate of 7.31 per cent between 1990 and 2010. The GDP growth rate declined from 8.15 per cent in 1997 to 5.76 per cent and 4.77 per cent respectively in 1998 and 1999, immediately after the 1997 Asian financial crisis. The period 2000-2007, which is the sample period of the empirical analyses of the thesis, gained the highest average annual rate of GDP growth in comparison to the other periods presented in Table 2.5. Since 2008, the economy has experienced a slowdown in GDP growth during the global economic crisis. In particular, the GDP growth rate declined to 6.31 per cent and 5.32 per cent respectively in 2008 and 2009. The economy appeared to slightly improve in 2010, achieving a GDP growth rate of 6.78 per cent which, however, fell far short of the growth rate of 8.46 per cent in 2007.

With respect to consumption, the private sector experienced positive growth during the entire sample while government expenditure maintained positive growth almost annually. There is some evidence indicating that government consumption grew higher than private consumption during the period 2008-2010, which could be the consequence of the stimulus packages in the context of economic slowdown.

On average, and compared with other periods, the period 2000-2007 experienced the highest rates of gross domestic savings and gross domestic capital formation. An improvement in domestic savings can be seen by comparing the gross domestic saving rate of 2.91 per cent of GDP in 1990 with those of 27.12 per cent and 26.96 per cent in 2000 and 2010, respectively, given that the economy maintained positive GDP growth during the entire period. In regard to the gross domestic capital formation, it reached the highest growth rate of 26.80 per cent in 2007 but declined sharply to 6.28 per cent and then 4.31 per cent, respectively, in 2008 and 2009. There was then an improvement in 2010, with a growth rate of 10.41 per cent, which however is much lower than the growth rate obtained in 2007.

From 1996 to 2007, the economy experienced a single-digit inflation rates ranging from 1.95 to 8.84. The global financial crisis of 2008 has worsened further the macro instability of the Vietnamese economy. The high rates of inflation could be in part due to the economic growth strategy which had been targeted at achieving the high rates of growth essentially by an over-reliance on the investment-led growth strategy rather than an improvement of macroeconomic stability.[8] Inflation reached the rate of 22.14 in 2008, declined to the single-digit rate of 6.34 in 2009 but then rose up to the rate of 11.86 in 2010, illustrating the macro instability of the economy during the period 2008-2010. Here, it is worth recalling the basic economic premise that macroeconomic stability is the key to sustainable development. And, stability is conventionally conceptualised in terms of low inflation, stable exchange rates, and low budgetary and current account deficits (Blanchard et al., 2010).

Government expenditure as a percentage of GDP, which is considered a measure of government size, reached its highest level in 2009 when the economy was experiencing a slowdown. The size of government in 2010 remained larger than that in the pre-2008 period. Meanwhile, during the period 1990-2010, the government had rarely maintained a budget surplus. In particular, the budget deficit in 2009 was much higher than that experienced in any year, except for 1990.

With respect to other macroeconomic indicators, from 1995 to 2010, the money supply as a percentage of GDP grew almost annually. The growth of the money supply increased sharply from year to year, went up to 46.12 per cent in 2007 and remained high at 33.30 per cent in 2010. Regarding external trade, the economy experienced trade surpluses in 1999, 2000 and 2001 while experiencing a trade deficit in the other years of the period 1990-2010. Gross international reserves increased gradually from 1995 to 2008, but then decreased dramatically in 2009. External debt as a percentage of GNI decreased sharply from the

[8] According to a study of the Viet Nam Centre for Economics and Policy Research cited in the United Nations Country Team Statement for the Consultative Group Meeting of June 2011, the recurrent periods of high inflation are "largely" attributed to internal factors, including inefficient public investment paired with a large budget deficit, loose monetary policy leading to high money supply and credit growth, and lack of a coordination between monetary policy and fiscal policy.

http://www.un.org.vn/en/feature-articles-press-centre-submenu-252/1841-re-establishing-and-maintaining-macroeconomic-stability-and-protecting-the-poor.html average rate of 172.42 per cent in the period 1990-1999 to 37.81 per cent in the period 2000-2007 but then rose to 42.5 per cent in 2010. It is noteworthy that the external debt structure had been dominated by long-term debts. During the period 1990-2000, the economy had also witnessed an annual depreciation of the domestic currency, i.e. the Vietnam Dong.

Table 2.5: Several Indicators of the Macro Economy (1990-2010)

Year	Real GDP Growth (%)	Private Consum -ption Growth (%)	Govern-ment Consum -ption (%)	Gross Domes-tic Savings (% of GDP)	Gross Domes-tic Capital Form-ation Growth (%)	Govern-ment Expendi -ture (% of GDP)	Budget Balance (% of GDP)	External Trade Balance (% of GDP)
1990	5.09	...	...	2.91	...	21.89	-7.23	-0.63
1991	5.81	...	...	10.10	...	14.16	-0.66	-0.82
1992	8.70	...	...	13.77	...	19.82	-0.80	-0.61
1993	8.08	...	...	16.78	...	25.17	-3.42	-4.15
1994	8.83	...	...	17.08	...	25.01	-2.23	-7.31
1995	9.54	7.21	8.40	18.20	17.07	23.85	-1.25	-11.31
1996	9.34	9.09	7.43	17.21	14.23	23.12	-0.89	-11.25
1997	8.15	5.91	4.00	20.10	9.38	22.56	-3.91	-4.65
1998	5.76	4.47	3.22	21.49	12.63	20.34	-1.62	-3.63
1999	4.77	2.60	-5.70	24.57	1.20	21.21	-3.35	3.39
2000	6.79	3.08	5.01	27.12	10.11	22.59	-4.28	1.21
2001	6.89	4.47	6.60	28.82	10.77	24.37	-3.55	1.47
2002	7.08	7.63	5.38	28.67	12.73	24.16	-2.31	-3.01
2003	7.34	8.01	7.19	27.42	11.86	26.43	-2.18	-6.53
2004	7.79	7.09	7.77	28.53	10.54	26.19	0.16	-4.92
2005	8.44	7.26	8.20	30.32	11.15	27.30	-1.08	-4.61
2006	8.23	8.35	8.50	30.62	11.83	27.55	1.33	-4.56
2007	8.46	10.80	8.90	29.19	26.80	29.41	-0.98	-14.59
2008	6.31	9.34	7.52	26.58	6.28	27.73	0.67	-14.03
2009	5.32	3.13	7.60	27.23	4.31	31.80	-6.56	-8.55
2010	6.78	10.00	12.28	26.96	10.41	30.66	...	-6.67
Average 90-10	7.31	6.78	6.39	22.56	11.33	24.54	-2.21	-5.04
Average 90-99	7.41	5.86	3.47	16.22	10.90	21.71	-2.54	-4.10
Average 00-07	7.63	7.09	7.19	28.84	13.22	26.00	-1.61	-4.44
Average 08-10	6.14	7.49	9.13	26.92	7.00	30.06	-2.95	-9.75

Source: ADB website. Average values are calculated by using ADB data.

Table 2.5 (cont.): Several Indicators of the Macro Economy (1990-2010)

Year	GDP Deflator (%)	M2 Growth (%)	M2 (% of GDP)	External Debt (% of GNI)	Short term Debt (%)	Debt Service (%)	International Reserves (Mill. USD)	Exchange Rate (VND/ 1 USD)
1990	42.10	53.09	27.07	384.01	7.65	...	0.01	6483
1991	72.79	78.73	26.47	257.06	8.26	...	0.01	10037
1992	32.56	33.71	24.56	255.32	10.62	...	0.01	11202
1993	17.41	18.95	23.02	190.62	10.22	...	6.76	10641
1994	16.96	33.19	24.09	156.22	10.74	...	16.11	10966
1995	17.04	22.57	23.03	123.98	12.87	...	1379.09	11038
1996	8.70	22.70	23.78	108.19	14.30	4.11	1813.77	11033
1997	6.60	26.10	26.01	82.64	10.75	7.71	2098.12	11683
1998	8.84	25.57	28.37	84.09	9.77	9.05	2100.51	13268
1999	5.74	39.28	35.67	82.01	10.24	9.96	3423.44	13943
2000	3.41	56.25	50.47	41.73	7.20	7.49	3509.63	14168
2001	1.95	25.53	58.13	39.02	6.22	6.71	3765.13	14725
2002	3.96	17.65	61.44	38.63	5.89	6.04	4231.82	15280
2003	6.67	24.94	67.04	41.75	8.07	3.43	6359.13	15510
2004	8.18	29.45	74.42	40.38	11.90	2.70	7186.07	15746
2005	8.19	29.74	82.30	36.71	13.52	2.64	9216.47	15859
2006	7.27	33.59	94.70	31.23	13.04	2.13	13591.01	15994
2007	8.24	46.12	117.88	33.01	20.36	2.25	23747.75	16105
2008	22.14	20.31	109.23	28.48	17.14	1.88	24175.87	16302
2009	6.03	28.99	126.17	32.35	18.09	1.79	16803.13	17065
2010	11.86	33.30	140.80	...	...	...	...	18621
Average 90-10	15.08	33.32	59.27	104.37	11.34	4.85	NA	NA
Average 90-99	22.87	35.39	26.21	172.41	10.54	7.71	NA	NA
Average 00-07	5.98	32.91	75.80	37.81	10.78	4.17	NA	NA
Average 08-10	13.34	27.53	125.40	30.42	17.62	1.84	NA	NA

Source: ADB Website. Average values are calculated by using ADB data.

2.2.3 Growth and Diversity of External Trade

From the late 1970s to the early 1990s, Vietnam was a member of the Council for Mutual Economic Assistance (COMECON). On the other side, during this period, the United States imposed economic sanctions on the Vietnamese economy. In part due to these factors, Vietnam's external trade was heavily dependent on trade with COMECON member countries. By the early 1990s, following the dissolution of the COMECON and then the loss of traditional trading partners, the Vietnamese economy witnessed a significant decline in external trade. Table 2.6 shows the negative growth rates of exports and imports in 1991.

The opening-up policy proposed in the 1986 Renovation appeared not enough to integrate the Vietnamese economy into the global economy. By the mid-1990s, the economic sanctions imposed by the United States on Vietnam since the mid-1970s were lifted. The normalisation of diplomatic relationship between Vietnam and the United States was completed in 1995, paving the way for an explosion of external trade and inward foreign direct investment (FDI) in Vietnam (GSO, 20011a). Vietnam joined the Association of Southeast Asian Nations (ASEAN) in 1995, the Asia Pacific Economic Cooperation (APEC) in 1998 and the World Trade Organisation (WTO) in 2007. By 2010, Vietnam has established trade relations with 175 countries and territories, had economic agreements and bilateral trade agreements with more than 60 partners, and received foreign investment from more than 84 countries and territories (GSO, 2011a).

As can also be seen in Table 2.6, the ratio of exports plus imports to GDP, which is often used as a proxy for trade openness in the economic literature, increased from 54.26 per cent in 1991 to 142.40 per cent in 2010. The average growth rate of either exports or imports was about twice as high as that of GDP in the same period. The average values also show that Vietnam had experienced a deficit in its trade balance account over the period observed.

Table 2.6: Exports, Imports and Trade Balance (1995-2010)

Year	Exports plus Imports as % of GDP	Trade Balance (%)	Export Growth (%)	Exports as % of GDP	Import Growth (%)	Imports as % of GDP
1990	54.13	-0.63	23.54	26.75	7.26	27.38
1991	54.26	-0.82	-13.19	26.72	-15.04	27.54
1992	50.78	-0.61	23.66	25.08	8.68	25.69
1993	49.44	-4.15	15.67	22.65	54.43	26.80
1994	57.11	-7.31	35.80	24.90	48.47	32.21
1995	61.44	-11.31	34.42	25.07	39.98	36.38
1996	70.10	-11.25	33.16	29.42	36.64	40.68
1997	73.08	-4.65	26.59	34.22	4.03	38.86
1998	72.44	-3.63	1.91	34.40	-0.80	38.04
1999	77.08	3.39	23.30	40.23	2.11	36.84
2000	91.49	1.21	25.49	46.35	33.17	45.14
2001	90.48	1.47	3.77	45.98	3.72	44.50
2002	98.29	-3.01	11.16	47.64	21.75	50.65
2003	108.41	-6.53	20.61	50.94	27.91	57.47
2004	121.53	-4.92	31.44	58.30	26.58	63.23
2005	127.24	-4.61	22.51	61.32	14.99	65.93
2006	135.32	-4.56	22.74	65.38	22.12	69.94
2007	151.35	-14.59	21.93	68.38	39.82	82.97
2008	151.66	-14.03	29.08	68.81	28.60	82.85
2009	126.05	-8.55	-8.90	58.75	-13.30	67.30
2010	142.40	-6.67	26.40	67.86	21.20	74.53
Average 90-10	93.53	-5.04	19.58	44.25	19.63	49.28
Average 90-94	53.14	-2.70	17.09	25.22	20.76	27.92
Average 95-10	106.15	-5.76	20.35	50.19	19.28	55.96
Average 00-07	115.51	-4.44	19.96	55.54	23.76	59.98

Source: ADB Website. Average values are calculated by using ADB data.
Note: Exports, Imports, and Exports plus Imports as percentages of GDP are expressed in current market prices.

Table 2.7 presents the direction of exports and imports which had changed dramatically between 1991 and 2010. And, the most significant change in regard to the direction of external trade can be seen in the trade relations between Vietnam and the United States. In particular, the proportion of exports to the U.S. market in Vietnam's total exports increased from 0 per cent in 1991 to 25.30 per cent in 2010.

Table 2.7: Direction of External Trade (1990-2010)

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Exports to																					
U.S.	0.00	0.00	0.00	0.00	2.34	3.12	2.81	3.12	5.01	4.37	5.06	7.09	14.68	19.55	18.97	18.26	19.70	20.81	18.93	20.76	25.30
Japan	14.16	34.47	32.31	31.39	29.09	26.81	21.31	18.24	16.18	15.48	17.78	16.70	14.59	14.44	13.37	13.38	13.16	12.54	13.62	11.08	13.02
China	0.32	0.93	3.71	4.55	7.29	6.64	4.69	5.16	4.70	6.47	10.61	9.43	9.09	9.35	10.95	9.95	8.14	7.51	7.24	7.08	11.10
Australia	0.32	0.25	0.83	1.83	1.14	1.02	0.89	2.51	5.04	7.06	8.79	6.93	7.95	7.05	7.12	8.39	9.40	7.83	6.74	4.29	5.02
Singapore	8.09	20.37	15.56	12.74	14.64	12.66	17.78	13.24	7.92	7.59	6.12	6.94	5.75	5.09	5.61	5.91	4.55	4.60	4.24	3.61	2.55
Germany	1.72	0.32	1.33	1.68	2.84	4.00	3.14	4.48	5.90	5.67	5.04	4.80	4.36	4.24	4.02	3.35	3.63	3.82	3.31	4.13	5.03
Malaysia	0.21	0.69	2.65	1.87	1.60	2.03	1.07	1.54	1.23	2.22	2.86	2.24	2.08	2.25	2.36	3.17	3.15	3.20	3.12	2.99	3.03
U.K.	0.08	0.11	1.07	0.77	1.37	1.37	1.72	2.89	3.59	3.65	3.31	3.40	3.42	3.75	3.81	3.13	2.96	2.95	2.52	2.53	3.72
Philippines	2.37	0.04	0.04	0.05	0.09	0.76	1.82	2.62	4.28	3.41	3.30	2.45	1.89	1.69	1.88	2.55	1.97	1.99	2.91	3.05	2.49
Korea	1.11	2.46	3.62	3.33	2.13	4.32	7.69	4.54	2.45	2.77	2.43	2.70	2.81	2.44	2.30	2.05	2.12	2.56	2.85	2.76	2.88
Imports from																					
China	0.17	0.79	1.25	2.18	2.48	4.04	2.95	3.49	4.48	5.73	8.96	9.90	10.93	12.43	14.37	16.05	16.46	20.25	19.39	18.87	36.36
Singapore	18.06	30.89	32.33	26.97	19.67	17.48	18.24	18.36	17.08	16.00	17.23	15.28	12.83	11.39	11.32	12.19	13.98	12.13	11.64	10.99	11.62
Korea	1.93	6.50	8.31	12.27	12.37	15.37	15.99	13.50	12.36	12.65	11.21	11.63	11.54	10.40	10.51	9.78	8.71	8.51	8.75	8.92	12.86
Japan	6.14	6.74	9.42	11.53	10.05	11.23	11.31	13.02	12.88	13.78	14.72	13.46	12.68	11.81	11.11	11.08	10.47	9.86	...	10.25	10.25
Thailand	0.62	0.61	1.62	2.54	3.87	5.39	4.44	4.96	5.86	4.78	5.19	4.89	4.84	5.08	5.81	6.46	6.76	5.97	6.08	7.34	9.19
Hong Kong	7.15	8.33	5.62	3.71	5.47	5.14	6.81	5.17	4.85	4.30	3.82	3.31	4.08	3.92	3.36	3.36	3.21	3.11	3.26	5.10	5.83
Malaysia	0.03	0.26	1.41	0.63	1.13	2.34	1.80	1.96	2.16	2.60	2.49	2.86	3.46	3.66	3.80	3.42	3.30	3.65	3.22	3.09	6.82
U.S.	0.02	0.04	0.08	0.10	0.76	1.60	2.21	2.17	2.84	2.75	2.33	2.54	2.32	4.53	3.55	2.35	2.20	2.71	3.26	4.89	3.68
India	0.16	1.02	0.35	0.21	0.48	0.75	0.79	0.73	0.94	1.17	1.14	1.41	1.64	1.81	1.86	1.62	1.96	2.16	2.59	2.64	2.67
Switzerland	0.47	0.21	1.55	0.34	0.45	0.92	1.27	1.25	0.82	0.87	0.75	0.71	0.70	1.17	2.15	2.43	3.02	1.62	2.34	2.09	3.06

Source: Calculation using Data from ADB Website

Foreign-invested enterprises have had a significantly increasing role in the growth of external trade, illustrated by their increased shares in exports from 27.04 per cent in 1995 to 54.14 per cent in 2010 and imports from 30.7 per cent in 2001 to 43.6 per cent in 2010, as respectively shown in Tables 2.8 and 2.10.

Table 2.8: Growth and Composition of Exports by Economic Agents (1995-2010)

Year	Exports by Economic Agents (%)			
	Domestic		Foreign	
	Export Share	Export Growth	Export Share	Export Growth
1995	72.96	...	27.04	...
1996	70.30	28.30	29.70	46.29
1997	65.02	17.08	34.98	49.10
1998	65.65	2.90	34.35	0.06
1999	59.44	11.62	40.56	45.63
2000	52.98	11.85	47.02	45.46
2001	54.77	7.28	45.23	-0.18
2002	52.88	7.33	47.12	15.79
2003	49.57	13.06	50.43	29.08
2004	45.30	20.12	54.70	42.58
2005	42.82	15.80	57.18	28.07
2006	42.10	20.67	57.90	24.29
2007	42.81	23.99	57.19	20.44
2008	44.93	35.48	55.07	24.30
2009	46.81	-5.11	53.19	-12.02
2010	45.86	23.88	54.14	28.69
Average 95-10	53.39	15.62	46.61	25.84
Average 95-99	66.67	14.97	33.33	35.27
Average 00-07	47.90	15.01	52.10	25.69
Average 08-10	45.86	18.08	54.14	13.66

Source: Calculation using Data from GSO (2004, 2011a)

Export composition, which is presented in Table 2.9, has undergone a structural transformation, witnessing an increased share of manufactured commodities. During the early years of the reform period, the composition of exports was dominated by agricultural commodities, and this can be attributed to agricultural reforms, especially the price mechanism and the autonomy for farmers, which promoted agricultural production. As the economic reforms expanded to trade and investment, the transformation process witnessed an expansion of the labour-intensive manufactured exports and, more recently, a growing share of component production and assembly. However, the export composition remains dominated by natural resources-based and labour-intensive products.

Table 2.9: Merchandise Exports by Economic Sectors and Commodities (1995-2010)

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
By Economic Sectors																
Heavy Industries	25.28	28.74	28.02	27.87	31.27	37.16	34.91	31.75	32.19	36.41	36.06	36.23	34.28	37.03	30.86	27.84
Light Industries	28.44	28.96	36.72	36.62	36.77	33.85	35.72	40.62	42.67	41.05	40.95	41.13	42.61	39.72	44.80	45.05
Agriculture	46.27	42.29	35.26	35.51	31.96	28.98	29.37	27.63	25.15	22.55	22.97	22.62	23.07	22.68	22.89	23.29
Others	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.03	0.58	1.44	3.81
By Commodities (SITC Rev.1)																
Food and live animals	37.90	33.40	29.30	33.70	28.50	26.10	27.00	24.60	22.00	19.90	19.60	18.90	18.90	19.40	20.20	...
Beverages and tobacco	0.10	0.10	0.40	0.10	0.10	0.10	0.30	0.50	0.80	0.70	0.50	0.40	0.30	0.30	0.40	...
Crude materials, inedible, except fuels	6.80	6.90	4.10	3.00	2.60	2.70	2.70	3.10	3.10	3.10	3.80	4.60	4.50	4.00	3.40	...
Mineral fuels, lubricants and related materials	22.20	21.70	18.00	16.50	20.60	26.40	23.10	21.40	20.60	23.50	25.80	24.40	20.70	20.30	14.90	...
Animal and vegetable oils and fats	0.30	0.50	0.30	0.20	0.20	0.50	0.20	0.10	0.10	0.10	0.10	0.00	0.10	0.20	0.10	...
Chemicals	0.60	0.90	1.20	1.00	1.30	1.10	1.50	1.60	1.70	1.60	1.70	2.00	2.10	2.30	2.20	...
Manuf. goods classified chiefly by material	6.40	5.30	6.10	4.70	7.50	6.30	6.60	6.70	6.70	7.10	6.70	7.30	8.20	10.20	9.20	...
Machinery and transport equipment	1.60	5.70	8.20	8.60	8.50	8.80	9.30	8.00	8.90	9.70	9.70	10.50	11.50	11.80	13.00	...
Miscellaneous manufactured articles	24.10	25.50	32.40	32.10	30.80	28.00	29.30	34.10	36.00	34.20	32.30	31.80	33.50	31.00	35.20	...
Commodities not elsewhere specified	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.60	1.40	...

Source: Calculation using Data from GSO Website

With respect to the composition of imports, Table 2.10 shows that the majority of imports have been production inputs which, as GSO (2011a) notes, essentially include machines, equipment and other inputs such as petrol, steel, plastics, chemicals, and intermediate goods of clothing and footwear industries. Also according to GSO (2011a), there are three factors explaining the high growth of imports. The first factor is the growth of demand on parts and other intermediate inputs to produce both manufactured exports and goods distributed in the domestic market. In addition, the process of industrialisation requires modern machines and technologies which the economy, dominated by low-tech industries of Vietnam, has been unable to produce. Furthermore, increased prices of some imported production inputs contributed to an increase in the overall values of imports.

Table 2.10: Composition of Imports by Economic Agents and Commodities (2001-2010)

Year	Imports by Economic Agents (%)		Imports by Commodities (%)	
	Domestic	Foreign	Production Inputs	Consumption Goods and Gold
2001	69.3	30.7	92.1	7.9
2002	66.1	33.9	92.1	7.9
2003	65.1	34.9	92.2	7.8
2004	65.3	34.7	93.3	6.7
2005	62.9	37.1	89.6	10.4
2006	63.3	36.7	88.0	12.0
2007	65.4	34.6	90.5	9.5
2008	65.5	34.5	88.9	11.1
2009	62.7	37.3	90.2	9.8
2010	56.4	43.6	90.0	10.0
Average 01-10	32.8	67.2	90.7	9.3

Source: GSO (2011a)

2.2.4 Sectoral GDP Growth and Structural Transformation

As can be seen in Table 2.11, during the period 1986-2010, growth can be observed in almost all sectors of the economy. This period had also witnessed a transformation in ownership structure and sectoral structure of GDP.

From a sectoral perspective, industry was the leading sector, growing at an average rate of 9.06 per cent per year. Before 2008, this sector had achieved a growth rate that was higher than the overall GDP growth. Although agriculture achieved the lowest growth rate among the three sectors, this sector maintained positive growth since 1988, in response to agriculture reforms. Economic reforms have also promoted the expansion of the service sector, although over the periods 1986-1999 and 2000-2007, average growth rates of this sector were relatively lower than those of the industry sector. During the period 2008-2010, the service sector grew faster than the agriculture and industry sectors.

The differential growth performance across sectors has contributed to accelerating the structural transformation of the economy. Although the agriculture sector grew positively during this period, it had been gradually losing its share in total GDP, from 34.74 per cent of GDP in 1986 to 16.43 per cent in 2010. During the same period, the industry sector had played an increasingly important role in the economy by sustaining annual growth rates relatively higher in comparison to those of other sectors and expanding its GDP share from 26.82 per cent in 1986 to 41.94 per cent in 2010. Also during that period, the share of services in total GDP had remained rather stable and, hence, the structural transformation appears to occur mainly between the agriculture and industry sectors. This trend seems to be in line with what is expected from the early stage of industrialisation where the economy is moving away from primary production towards production concentrated in secondary and then tertiary sectors.

Another key structural transformation has been associated with the increasing role of the private sector in the economy. The GDP share of 42.65 per cent in 1986 of the state sector declined slightly to 36.99 per cent in 2010. The private sector accounted for an average 59.68

per cent per year in total GDP in the period 1986-2010. In terms of GDP contribution, the role of the foreign-invested sector was still less significant than that of the non-state sector. Specifically, during the period 1995-2010, foreign-invested enterprises contributed an average of 10.50 per cent per year to total GDP while the average contribution of the non-state sector was 49.20 per cent per year.

Table 2.11: Growth and Composition of GDP (1986-2010)

Year	Real GDP Growth (%)	GDP pc Growth (%)	GDP by Economic Agents (%)								GDP by Economic Sectors (%)					
			State		Non-State and Foreign-Invested						Agriculture		Industry		Services	
					Total		Non-State		Foreign-Invested							
			GDP Share	GDP Growth	GDP Share	GDP Growth	GDP Share	GDP Growth	GDP Share	GDP Growth	GDP Share	GDP Growth	GDP Share	GDP Growth	GDP Share	GDP Growth
1986	...	...	42.65		57.35	...	...	...	...	...	34.74	...	26.82	...	38.44	...
1987	3.63	1.40	43.51	5.71	56.49	2.08	...	...	...	...	33.14	-1.14	28.07	8.46	38.79	4.57
1988	6.01	3.89	44.18	7.65	55.82	4.76	...	...	...	...	32.40	3.65	27.80	5.00	39.80	8.77
1989	4.68	2.99	41.46	-1.76	58.54	9.77	...	...	...	...	33.12	7.00	25.87	-2.59	41.01	7.86
1990	5.09	3.12	38.06	-3.53	61.94	11.20	...	...	...	...	31.83	1.00	25.17	2.27	43.00	10.19
1991	5.81	3.88	38.35	6.63	61.65	5.31	...	...	...	...	30.74	2.18	25.63	7.71	43.64	7.38
1992	8.70	6.78	39.02	10.60	60.98	7.52	...	...	...	...	30.22	6.88	26.59	12.79	43.19	7.58
1993	8.08	6.22	39.55	9.54	60.45	7.14	...	...	...	...	28.88	3.28	27.71	12.62	43.41	8.64
1994	8.83	7.02	40.12	10.39	59.88	7.82	53.47	...	6.41	...	27.43	3.37	28.87	13.39	43.70	9.56
1995	9.54	7.76	40.07	9.42	59.93	9.62	53.20	8.98	6.73	14.98	26.24	4.80	29.94	13.60	43.82	9.83
1996	9.34	7.60	40.78	11.27	59.22	8.04	51.87	6.60	7.35	19.41	25.06	4.40	31.34	14.46	43.60	8.80
1997	8.15	6.48	41.35	9.67	58.65	7.11	50.44	5.18	8.20	20.76	24.17	4.33	32.64	12.62	43.20	7.14
1998	5.76	4.15	41.27	5.56	58.73	5.91	49.49	3.77	9.24	19.10	23.66	3.53	33.43	8.33	42.91	5.08
1999	4.77	3.21	40.40	2.55	59.60	6.33	49.24	4.24	10.36	17.56	23.76	5.24	34.36	7.68	41.88	2.25
2000	6.79	5.36	40.75	7.72	59.25	-0.39	48.43	5.04	7.16	-26.21	23.28	4.63	35.41	10.07	41.30	5.32
2001	6.89	5.56	40.96	7.44	59.04	13.52	48.19	6.36	10.85	7.21	22.43	2.98	36.57	10.39	41.00	6.10
2002	7.08	5.85	40.97	7.11	59.03	7.06	48.17	7.04	10.86	7.16	21.82	4.17	37.39	9.48	40.79	6.54
2003	7.34	6.10	41.09	7.65	58.91	7.12	47.73	6.36	11.18	10.52	21.06	3.62	38.48	10.48	40.45	6.45
2004	7.79	6.51	41.07	7.75	58.93	7.83	47.36	6.95	11.56	11.51	20.39	4.36	39.35	10.22	40.25	7.26
2005	8.44	7.18	40.67	7.37	59.33	9.18	47.26	8.21	12.07	13.22	19.56	4.02	40.17	10.69	40.27	8.48
2006	8.23	7.03	39.89	6.17	60.11	9.65	47.35	8.44	12.75	14.33	18.74	3.69	40.97	10.38	40.29	8.29
2007	8.46	7.29	38.96	5.91	61.04	10.14	47.75	9.37	13.29	13.04	17.93	3.76	41.63	10.22	40.44	8.85
2008	6.31	5.19	38.24	4.36	61.76	7.55	48.27	7.47	13.48	7.85	17.65	4.68	41.50	5.98	40.84	7.37
2009	5.32	4.21	37.76	3.99	62.24	6.15	48.82	6.52	13.42	4.81	17.07	1.82	41.58	5.52	41.35	6.63
2010	6.78	5.67	36.99	4.62	63.01	8.10	49.42	8.09	13.59	8.12	16.43	2.78	41.94	7.70	41.63	7.52
Average 86-10	6.99	5.44	40.32	6.41	59.68	7.44	49.20	6.79	10.50	10.21	24.87	3.71	33.57	9.06	41.56	7.35

Average 86-99	6.80	4.96	40.77	6.44	59.23	7.12	51.29	5.75	8.05	18.36	28.96	3.73	28.87	8.95	42.17	7.51
Average 00-07	7.63	6.36	40.55	7.14	59.46	8.01	47.78	7.22	11.22	6.35	20.65	3.90	38.75	10.24	40.60	7.16
Average 08-10	6.14	5.02	37.66	4.32	62.34	7.27	48.84	7.36	13.50	6.93	17.05	3.09	41.67	6.40	41.27	7.17

Source: Calculation using Data from GSO (2004, 2011a)

Note: GDP is expressed in the constant 1994 prices. The agriculture sector includes agriculture, forestry and fishery. The industry sector includes mining and quarrying; manufacturing; electricity, gas and water; and construction. The service sector includes trade; hotel and restaurant; transport and telecommunication; finance; real estates; and other activities.

As presented in Table 2.12, further decomposition of GDP by economic agents within the non-state sector reveals that households assumed a dominant role. Nevertheless, the influence of collectives and households within the non-state sector fell slightly from the mid-1990s to 2010, which is contrary to the gradually increasing role of private enterprises.

Table 2.12: Composition of GDP within the Non-State Sector (1994-2010)

Year	Share in GDP (%)			Share in Non-State GDP (%)		
	Collectives	Private Enterprises	Households	Collectives	Private Enterprises	Households
1994	10.17	7.44	35.86	16.99	12.42	70.59
1995	9.70	7.56	35.94	16.19	12.61	71.20
1996	9.19	7.65	35.03	15.52	12.91	71.57
1997	8.72	7.50	34.22	14.87	12.80	72.33
1998	8.54	7.50	33.45	14.54	12.78	72.69
1999	8.64	7.51	33.09	14.50	12.60	72.90
2000	8.53	7.72	32.18	15.35	13.88	70.77
2001	...	...	...	...	...	...
2002	...	...	...	...	...	...
2003	...	...	...	...	...	...
2004	...	...	...	...	...	...
2005	7.19	9.71	30.36	12.11	16.37	71.52
2006	...	...	...	...	...	...
2007	6.55	11.00	30.21	10.72	18.01	71.26
2008	6.34	11.48	30.45	10.27	18.58	71.14
2009	6.19	11.93	30.70	9.95	19.16	70.89
2010	5.97	12.39	31.06	9.48	19.67	70.85
Average 94-10	7.98	9.11	32.71	13.37	15.15	71.48
Average 94-99	9.16	7.53	34.60	15.44	12.69	71.88
Average 00-07	7.42	9.47	30.92	12.73	16.09	71.19
Average 08-10	6.17	11.93	30.74	9.90	19.14	70.96

Source: Calculation using Data from GSO (2004) and GSO Website
Note: GDP is expressed in the constant 1994 prices.

Redefining the role of the state sector and the private sector in the economy has not achieved impressive results if it is observed from the overall GDP-ownership structure. More precisely, the state sector still maintained a very large share of total GDP, although the GDP contribution of the private sector had been increasing over time. The process of structural transformation in terms of ownership was likely to be faster in the industry sector. Table 2.13 shows that, the industry GDP share of the state sector declined from 41.80 per cent in 2000 to 22.11 per cent in 2010. Meanwhile, growth rates of industry GDP contributed by the non-state and foreign-invested sectors were much higher than that of the state sector, indicating that the private sector has played an increasingly important role as the engine of industrial development.

Table 2.13: Growth and Composition of Industry GDP by Economic Agents (2000-2010)

Year	Overall Industry Growth (%)	Structure of Industry GDP by Ownership (%)					
		State		Non-State		Foreign-Invested	
		Industry GDP Share	Industry GDP Growth	Industry GDP Share	Industry GDP Growth	Industry GDP Share	Industry GDP Growth
2000	...	41.80	...	22.26	...	35.94	...
2001	14.63	41.10	12.71	23.60	21.53	35.30	12.59
2002	14.85	40.26	12.51	24.31	18.32	35.43	15.25
2003	16.85	38.56	11.91	25.66	23.34	35.78	18.00
2004	16.57	37.02	11.92	26.93	22.34	36.04	17.44
2005	17.15	33.87	7.19	28.85	25.47	37.28	21.17
2006	16.81	30.69	5.82	31.05	25.73	38.26	19.88
2007	16.75	27.60	4.99	33.17	24.71	39.23	19.71
2008	13.92	24.88	2.71	34.86	19.75	40.25	16.88
2009	7.79	23.50	1.80	35.64	10.19	40.86	9.42
2010	13.96	22.11	7.24	35.87	14.70	42.02	17.19
Average 00-07	16.23	36.36	9.58	26.98	23.06	36.66	17.72
Average 08-10	11.89	23.50	3.92	35.46	14.88	41.05	14.50
Average 00-10	14.93	32.85	7.88	29.29	20.61	37.86	16.75

Source: Calculation using Data from GSO (2011a)
Note: GDP is expressed in the constant 1994 prices. The industry sector reported in this table does not include the construction sector.

As can be seen in Table 2.14, the high ratio of investment to GDP increased from 33.08 per cent in 1995 to 72.55 per cent in 2010, which indicates that physical capital accumulation has been an important factor determining GDP growth. From the sectoral perspective, the importance of the industry sector in the economy raises a question in regard to the distribution of production inputs among economic sectors. Table 2.14 also shows that the share of the industry sector in total investment increased from 26.99 per cent in 1987 to 41.99 per cent in 2009. Meanwhile, the investment share of the agriculture sector decreased from 13.27 per cent to 6.89 per cent, and that of the service sector decreased slightly from 59.74 per cent to 51.12 per cent. The employment issue is discussed in the next section on the evolution of the labour market.

Table 2.14: Ratio and Composition of Investment (1986-2010)

Year	Ratio of Invest-ment to GDP (%)	Composition of Investment (%)					
		By Economic Agents			By Economic Sectors		
		State	Non-State	Foreign-Invested	Agricul-ture	Industry	Services
1986	...	59.75	40.25	0.00	...	...	...
1987	...	52.82	47.18	0.00	13.27	26.99	59.74
1988	...	53.32	44.16	2.52	12.99	31.00	56.01
1989	...	45.54	40.88	13.58	9.40	33.32	57.27
1990	...	40.20	46.75	13.06	11.15	26.61	62.25
1991	...	37.97	47.73	14.30	12.42	30.86	56.72
1992	...	35.12	43.92	20.96	9.74	29.91	60.35
1993	...	44.00	30.82	25.18	7.29	49.04	43.67
1994	...	38.30	31.31	30.39	7.56	44.55	47.89
1995	33.08	42.03	27.61	30.37	13.27	34.07	52.66
1996	34.75	49.08	24.94	25.97	13.04	35.99	50.97
1997	38.31	49.43	22.61	27.96	13.10	33.87	53.03
1998	37.18	55.52	23.73	20.75	12.78	35.57	51.65
1999	38.96	58.67	24.05	17.28	14.15	36.98	48.87
2000	42.06	59.15	22.88	17.97	13.81	39.27	46.92
2001	44.25	59.80	22.59	17.61	9.50	42.41	48.09
2002	47.24	58.57	23.74	17.69	8.72	42.43	48.85
2003	49.61	57.23	25.68	17.08	8.45	41.54	50.00
2004	52.24	55.51	28.28	16.22	7.76	42.47	49.77
2005	54.43	53.85	29.37	16.78	7.48	42.30	50.22
2006	57.20	52.03	29.96	18.00	7.56	42.99	49.45
2007	67.00	42.67	29.93	27.40	6.73	42.83	50.44
2008	67.94	38.59	26.81	34.60	7.11	41.14	51.74
2009	71.88	46.62	24.99	28.39	6.89	41.99	51.12
2010	72.55	44.93	26.88	28.19	...	...	...
Average 86-10	50.54	49.23	31.48	20.97	10.50	37.32	52.19
Average 86-99	36.46	47.27	35.42	20.19	11.99	34.02	53.99
Average 00-07	51.76	54.85	26.55	18.59	8.75	42.03	49.22
Average 08-10	70.79	43.38	26.23	30.39	7.00	41.57	51.43

Source: Calculation using Data from GSO (2004, 2011a).

Note: From 1986 to 1994, investment is expressed in the constant 1982 prices. From 1995 to 2010, investment is expressed in the constant 1994 prices. The agriculture sector includes agriculture, forestry and fishery. The industry sector includes mining and quarrying; manufacturing; electricity, gas and water; and construction. The service sector includes trade; hotel and restaurant; transport and telecommunication; finance; real estates; and other activities.

2.2.5 Evolution of the Labour Market

The population of Vietnam reached 86.92 million by 2010. As can be seen in Table 2.15, the population grew annually at an average rate of 1.06 per cent during the period 1986-2010. The proportion of urban population in total population, which is referred to as an indicator of urbanisation rate, increased from 19.34 per cent in 1986 to 30.17 per cent in 2010. According to GSO (2011b), the patterns of urbanisation of the last three decades indicate that the process of urbanisation has been closely tied to the process of migration. In particular, migrants tend to live in urban areas and therefore contribute substantially to the increased urban population, tend to work as wage employees, and tend to have moved for job related reasons.

According to GSO (2011c), the total dependency ratio has declined over time, reaching the level of a “golden population structure” by the end of 2007, and fell further to 44.7 in 2009. In addition, the period of “golden population structure” is predicted to continue over the next three decades. At the province level, in 2009, the “golden population structure” could be seen in 43 of 63 provinces. GSO (2011c) also points out that the relatively low quality of the labour force remains a major constraint for the country to take full advantage of its “golden population structure” and, therefore, suggests that it is necessary to enhance the labour quality through healthcare, education and high-tech training programme as well as to increase the number of job opportunities, especially jobs requiring highly-qualified and highly-productive labour.

The rate of labour participation in Vietnam is considered to be high (Pierre, 2012). The Labour Force Surveys conducted by the General Statistics Office of Vietnam in 2007 and 2009 show an increase in labour participation rate. In particular, the most significant increase can be seen from the 15-19 year-old group and the over 50 year-old group. The rate increased from 37.3 per cent in 2007 to 43.8 per cent in 2009 for the former and from 55.6 per cent to 58.9 per cent over the same period for the latter. The labour participation rate is higher in the urban areas but lower in the rural areas, and this could be in part due to the fact that the proportion of wage employment in total employment is higher in the former compared with that in the latter. Table 2.15 also shows an annually increased employment-to-population ratio. According to Pierre (2012), the increases in both the labour participation rate and the employment-to-population ratio suggest that there was a net increase in the number of people participating into the labour force and that the number of jobs created nearly kept up with the net influx of labour participants, although there could exist adjustments particularly in the employment type and the employment quality.

In regard to labour quality, the labour force was dominated by unskilled workers, who had only primary education. According to the Ministry of Labour, Invalids and Social Affairs of Vietnam (cited in Zhu, 2011), the proportion of unskilled labour in the workforce slightly decreased from 49 per cent in 1993 to 44 per cent in 2006, and semi-skilled labour, those having secondary education, took up most of this decrease. The proportion of skilled labour, those having tertiary education, is still very small though increasing from 1.8 per cent to 4.2 per cent over the same period. As suggested by the patterns of labour force status by educational levels obtained from the 2007 and 2009 Labour Force Surveys, people with low levels of educational attainment are less likely to work as wage employees than those having higher education (Pierre, 2012).

With respect to the composition of employment, presented in Table 2.15, the employment share of the industry sector increased from 11.24 per cent in 1990 to 21.54 per cent in 2009 while that of the service sector grew from 15.74 per cent to 26.54 per cent. The agriculture sector still accounted for a very large share of employment, although there was a decline from 73.02 per cent in 1990 to 51.92 per cent in 2009. The distribution of investment and employment among economic agents over the period 1990-2010, as presented in Tables 2.14 and 2.15, shows that the state sector maintained a large share in total investment while the private sector has played an important role in job creation. Regarding the size of the informal sector in the labour market, as can be estimated from data collected by the Labour Force Surveys, the share of the informal sector in total employment decreased between 2007 and 2009 (Pierre, 2012). According to the 2007 Labour Force Survey, the shares of agricultural employment, informal employment, and formal employment in total employment are 50.1 per cent, 27.7 per cent, and 22.2 per cent, respectively. Those respective figures obtained from the 2009 Labour Force Survey are 47.6 per cent, 24.4 per cent, and 27.9 per cent, respectively. The increase in the proportion of people working in the formal sector can be seen in all age groups, which mostly took place in the groups of young people, i.e. the 15-24 year-old group and the 25-34 year-old group.

Table 2.15 shows a decrease in the overall unemployment rate from 4.50 per cent in 1999 to 2.39 per cent in 2009. Over the period 2000-2007, the overall unemployment was at an average rate of 2.57 per cent per year. The urban unemployment rate was much higher than the overall rate, suggesting that the pressure to generate new job opportunities should be much more intense in the urban areas where, compared with the rural areas, there are more people seeking for opportunities of wage employment. As can also be seen in Table 2.15, the overall underemployment rate was very high, ranging between 4.9 per cent and 14.4 per cent during the period 1996-2007, and the rural underemployment rate was higher than the urban underemployment rate.

Table 2.15: Population, Employment, Unemployment and Underemployment (1989-2010)

Year	Population			Employment								Unemployment		Underemployment		
						Composition (%)										
						By Economic Agent			By Economic Sector							
	Total (Thousand Persons)	Growth (%)	Ratio of Urban to Total Population (%)	Growth (%)	Ratio to Population (%)	State	Non-State	Foreign-Invested	Agricul-ture	Industry	Services	Overall (%)	Urban (%)	Overall (%)	Urban (%)	Rural (%)
1986	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1987	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1988	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1989	64774.0	1.64	19.68	...	...	...	...	...	...	...	...	...	...	...	...	...
1990	66016.7	1.92	19.51	...	...	11.61	...	...	73.02	11.24	15.74	...	...	...	...	...
1991	67242.4	1.86	19.67	2.46	44.55	10.41	...	...	72.70	11.25	16.05	...	...	...	...	...
1992	68450.1	1.80	19.85	2.39	44.81	9.64	...	...	72.40	11.26	16.34	...	...	...	...	...
1993	69644.5	1.74	20.05	2.34	45.08	9.37	...	...	72.06	11.28	16.66	...	...	...	...	...
1994	70824.5	1.69	20.37	2.29	45.34	9.06	...	...	71.68	11.31	17.01	...	...	...	...	...
1995	71995.5	1.65	20.75	2.25	45.61	9.24	...	...	71.25	11.37	17.38	...	5.88	...	...	...
1996	73156.7	1.61	21.08	2.21	45.88	9.29	...	...	70.72	11.52	17.77	...	6.01	10.2	9.2	10.4
1997	74306.9	1.57	22.66	2.17	46.15	9.47	...	...	70.15	11.66	18.20	...	6.85	17.3	9.3	19.5
1998	75456.3	1.55	23.15	2.14	46.42	9.60	...	...	69.55	11.80	18.65	...	6.74	13.1	8.1	11.6
1999	76596.7	1.51	23.61	2.11	46.69	9.54	...	...	68.91	11.95	19.13	4.50	6.42	10.6	8.1	11.6
2000	77630.9	1.35	24.12	3.06	46.97	11.75	87.28	0.97	67.55	11.99	20.46	4.40	6.42	8.6	6.6	9.1
2001	78620.5	1.27	24.55	2.98	47.76	11.72	87.37	0.91	...	...	...	2.34	6.28	14.4	8.4	16.2
2002	79537.7	1.17	24.99	2.87	48.56	11.80	87.12	1.08	...	...	...	2.53	6.01	13.7	8.6	15.2
2003	80467.4	1.17	25.76	2.87	49.38	12.17	85.96	1.86	...	...	...	2.23	5.78	11.8	7.7	13.1
2004	81436.4	1.20	26.53	2.91	50.21	12.10	85.70	2.20	...	...	...	2.17	5.60	9.2	5.7	10.3
2005	82392.1	1.17	27.10	2.88	51.06	11.61	85.79	2.60	57.10	18.20	24.70	2.12	5.31	8.1	4.5	9.3
2006	83311.2	1.12	27.66	2.82	51.92	11.18	85.82	3.01	55.37	19.23	25.40	2.52	4.82	...	...	...
2007	84218.5	1.09	28.20	2.79	52.79	11.03	85.51	3.46	53.91	19.98	26.12	2.26	4.64	4.9	2.1	5.8
2008	85118.7	1.07	28.99	2.77	53.68	10.89	85.46	3.65	49.25	20.83	29.92	2.00	4.65	...	...	...
2009	86025.0	1.06	29.74	2.76	54.58	10.56	86.25	3.19	51.92	21.54	26.54	2.39	4.60	...	...	...
2010	86927.7	1.05	30.17	2.73	55.50	10.41	86.07	3.52	...	...	...	...	4.29	...	...	...

Average 86-10	NA	1.48	23.48	2.59	48.65	10.59	86.21	2.40	65.47	14.15	20.38	2.68	5.64	...	...	...
Average 86-99	NA	1.75	20.66	2.26	45.62	9.73	...	...	71.24	11.46	17.29	4.50	6.38	...	...	...
Average 00-07	NA	1.19	26.11	2.90	49.83	11.67	86.32	2.01	58.48	17.35	24.17	2.57	5.61	10.10	6.23	11.29
Average 08-10	NA	1.06	29.63	2.76	54.59	10.62	85.93	3.45	50.58	21.19	28.23	2.19	4.51	...	...	...

Source: GSO (2004, 2011a), MOLISA Website, and ADB Website. Growth rates, ratios, and composition are calculated.

Note: The agriculture sector includes agriculture, forestry and fishery. The industry sector includes mining and quarrying; manufacturing; electricity, gas and water; and construction. The service sector includes trade; hotel and restaurant; transport and telecommunication; finance; real estates; and other activities.

2.2.6 Explosion of Private Entrepreneurship

A number of legal documents promulgated in the late 1990s and the 2000s, as presented previously, have played an important role in creating a level playing field for all types of enterprises. Consequently, the number of private sector enterprises increased from nearly zero in the late 1980s to some hundred thousand in 2007. The characteristics of Vietnamese enterprises, as well as the development of the private sector, are described in the next paragraphs using the aggregate data collected from the Annual Enterprise Surveys, which is published in GSO (2010).

Table 2.16 presents the average annual growth rate in the number of enterprises and the average annual distribution of enterprises by ownership and economic sectors of the period 2000-2007. The state sector accounted for the average share of 6.62 per cent per year in the total number of enterprises, which was much smaller than the average share of 89.89 per cent per year of the non-state sector but about twice the average annual share of 3.49 per cent per year of the foreign-invested sector. While the non-state sector and the foreign-invested sector achieved high average rates of annual growth of 22.88 per cent and 18.49 per cent, respectively, the state sector experienced a decline, i.e. an average growth rate of -6.83 per cent per year, in its number of enterprises.

Within the non-state sector, joint-stock companies without state capital had the highest average growth rate of 76.24 per cent per year, followed by collective-name enterprises with the average growth rate of 75.37 per cent per year. Nevertheless, both had very small average shares of the total number of enterprises, i.e. 6.72 per cent per year and 0.02 per cent per year, respectively, as well as in the total number of non-state enterprises, i.e. 7.30 per cent per year and 0.03 per cent per year, respectively. The non-state sector was dominated by private enterprises and private limited companies. On average per year, the former accounted for 35.68 per cent in the total number of enterprises and 40.14 per cent in the number of non-state enterprises. The figures for the latter are 40.61 per cent per year and 44.85 per cent per year, respectively. Concerning the growth rates, the latter grew annually at the average rate of 33.65 per cent, which is much faster than the rate of 10.25 per cent of the former. Collectives did not comprise a very large share in the number of enterprises in comparison to other types of enterprises, although they grew annually at the average rate of 11.34 per cent. It is worth noting that, although joint-stock companies with state capital accounted for a very small average share of the number of enterprises, their annual average growth rate of 27.22 per cent was relatively higher than that of the non-state sector.

The foreign-invested sector was dominated by 100 per cent foreign-owned enterprises which, on average, accounted for 71.26 per cent per year of the number of foreign-invested enterprises and grew annually at 25.20 per cent in terms of the number of enterprises. Also belonging to the foreign-invested sector, joint ventures, however, experienced a less impressive growth rate in comparison to other economic agents. It could be that, during this period, progress in trade and investment liberalisation had gradually removed barriers for FDI in some specific sectors and, as a result, foreign investors did not need to enter a joint-venture relationship with domestic enterprises. Therefore, the number of 100 per cent foreign-owned enterprises grew much faster than that of joint ventures. Another possibility could be that previous joint ventures became 100 per cent foreign-owned enterprises after the foreign investors bought the rest of the shares from domestic shareholders.

Growth and distribution of the number of enterprises can also be viewed from the sectoral perspective. In line with the previous section, economic sectors are divided into three main categorises: agriculture, industry and services. As can also be seen in Table 2.16, on average per year, the service sector comprised the largest share of enterprises of 58.78 per cent, followed by the industry sector with 37.28 per cent. The agriculture sector accounted for an average share of 3.94 per cent per year in the number of enterprises but experienced a negative annual average growth rate of -3.86 per cent per year. The growth rate of the number of enterprises belonging to the industry sector was relatively similar to that of the service sector. In terms of the number of enterprises, the growth rates of both sectors were slightly higher than the growth rate of the whole economy, implying that, on average, each sector contributed relatively equally to the overall explosion in the number of enterprises. Within the industry sector, while machinery enterprises had the small average share of 5.85 per cent per year of the number of industry enterprises, non-machinery manufacturing enterprises had the large average share of 54.42 per cent per year. The machinery and non-machinery manufacturing sectors together dominated the industry sector in terms of the number of enterprises, growing annually at the average rates of 16.65 per cent and 17.00 per cent, respectively. Likewise, the wholesale, the retail and the business service sectors together accounted for a large share in the number of service enterprises. The business service sector had the average share of 13.28 per cent per year in the number of service enterprises, which is not small compared with those of 31.43 per cent per year and 23.49 per cent per year, respectively, of the wholesale and retail sectors, although it grew annually at a faster average rate.

Table 2.16: Growth in Number of Enterprises and Distribution of Enterprises by Ownership and by Economic Sector (2000-2007, Average Per Year)

Distribution of Enterprises by Ownership					Distribution of Enterprises by Economic Sector				
Economic Agent	Share in Total	Share in NSS	Share in FIS	Growth	Economic Sector	Share in Total	Share in IND	Share in SER	Growth
The State Sector	6.62	NA	NA	-6.83	Agriculture	3.94	NA	NA	-3.86
The Non-State Sector	89.89	NA	NA	22.88	Industry	37.28	NA	NA	21.06
- Collectives	5.93	6.66	NA	11.34	- Machinery Manufacturing	2.18	5.85	NA	16.65
- Private Enterprises	35.68	40.14	NA	10.25	- Non-Machinery Manufacturing	20.26	54.42	NA	17.00
- Collective-Name Enterprises	0.02	0.03	NA	75.37	Services	58.78	NA	NA	22.12
- Private Limited Companies	40.61	44.85	NA	33.65	- Wholesale Trade	18.52	NA	31.43	25.72
- Joint-Stock Companies with State Capital	0.92	1.02	NA	27.22	- Retail Trade	13.74	NA	23.49	14.31
- Joint-Stock Companies without State Capital	6.72	7.30	NA	76.24	- Business Services	7.85	NA	13.28	32.81
The Foreign-Invested Sector	3.49	NA	NA	18.49					
- 100 per cent Foreign-Owned Enterprises	2.47	NA	71.26	25.20					
- Joint Ventures	1.02	NA	28.74	5.00					
Total	NA	NA	NA	20.54	Total	NA	NA	NA	20.54

Source: Calculation using Data from GSO (2010)

Note: NSS and FIS represent the non-state sector and the foreign-invested sector, respectively. The agriculture sector includes enterprises operating in agriculture, forestry and fishery which are coded as A and B, respectively. The industry sector, denoted as IND, includes enterprises operating in mining and quarrying; manufacturing; electricity, water and gas supply; and construction, which have the codes C, D, E and F, respectively. The machinery manufacturing sector includes enterprises operating in D29, D30, D31, D32, D33, D34 and D35 manufacturing industries; whereas, the non-machinery manufacturing sector includes enterprises operating in the remaining manufacturing industries. The service sector, denoted as SER, contains the remaining industries of the economy. Wholesale and retail have the codes G51 and G52, respectively. The business service sector includes enterprises operating in finance, insurance, real estate and other business services, i.e. the J-coded and L-coded categories.

Enterprises can also be divided into four main categorises according to their workforce size. In accordance with the definition stipulated in the Decree No. 90/2001/CP-NĐ of the Government of Vietnam, a micro enterprise has less than 10 employees; the workforce of a small enterprise ranges from 10 to 50; the workforce of a medium-sized enterprises ranges from more than 50 to less than 300 workers; and a large enterprise has more than 300 employees. The period 2000-2007 was dominated by micro and small enterprises. On average per year, 55.41 per cent and 35.61 per cent of the total number of enterprises were micro and small, respectively. In contrast, medium and large enterprises accounted for very small average shares of 3.31 per cent per year and 5.67 per cent per year, respectively.

On the distribution of enterprises by workforce-size and ownership, as presented in Table 2.17, 99.23 per cent of micro enterprises and 87.36 per cent of small enterprises belonged to the non-state sector. In the non-state sector itself, medium and large enterprises accounted for the very small average shares of 2.08 per cent per year and 2.14 per cent per year, respectively. With respect to the distribution of medium enterprises, the non-state sector also accounted for the largest average share, i.e. 58.82 per cent per year, followed by the state sector with 29.26 per cent per year. On average per year, only 11.92 per cent of medium enterprises and 16.62 per cent of large enterprises belonged to the foreign-invested sector. The state sector accounted for the largest average share, i.e. 58.82 per cent per year, in the number of large enterprises, followed by the non-state sector with the average share of 36.10 per cent per year. While the state sector was dominated by large enterprises, followed by small enterprises; the reverse composition could be seen in the foreign-invested sector.

From the sectoral perspective, Table 2.18 shows that all three economic sectors were dominated by micro and small enterprises. While the average share of 58.04 per cent per year in the number of small enterprises and the average share of 55.74 per cent per year in the number of large enterprises are belonged to the industry sector, the service sector had the average share of 75.34 per cent per year in the number of micro enterprises and 58.87 per cent per year in the number of medium enterprises. These figures could indicate that, on average, while the majority of small and large enterprises were specialised in the industry sector, the majority of micro and medium enterprises belonged to the service sector.

Table 2.17: Distribution of Enterprises by Workforce Size and Economic Agents (2000-2007, Average Per Year)

The State Sector (%)				The Non-State Sector (%)				The Foreign-Invested Sector (%)			
Micro	Small	Medium	Large	Micro	Small	Medium	Large	Micro	Small	Medium	Large
0.18	7.34	29.26	47.29	99.23	87.36	58.82	36.10	0.59	5.30	11.92	16.62
Micro Enterprises (%)			Small Enterprises (%)			Medium Enterprises (%)			Large Enterprises (%)		
State	Non-State	Foreign	State	Non-State	Foreign	State	Non-State	Foreign	State	Non-State	Foreign
1.40	61.17	9.47	38.24	34.61	53.90	15.57	2.08	11.19	44.79	2.14	25.44

Source: Calculation using Data from GSO (2010)

Table 2.18: Distribution of Enterprises by Workforce Size and Economic Sectors (2000-2007, Average Per Year)

The Agriculture Sector (%)				The Industry Sector (%)				The Service Sector (%)			
Micro	Small	Medium	Large	Micro	Small	Medium	Large	Micro	Small	Medium	Large
2.84	5.98	2.20	3.28	21.83	58.04	38.93	55.74	75.34	35.97	58.87	40.98
Micro Enterprises (%)			Small Enterprises (%)			Medium Enterprises (%)			Large Enterprises (%)		
Agriculture	Industry	Services	Agriculture	Industry	Services	Agriculture	Industry	Services	Agriculture	Industry	Services
36.23	32.63	22.52	56.39	55.44	26.21	2.05	3.49	13.73	5.32	8.44	6.58

Source: Calculation using Data from GSO (2010)

To get an insight into the results of economic reforms towards an increased role of the private sector in the Vietnamese economy, it is useful to consider not only the above-presented ownership structure of output, but also the allocation of production inputs, namely, labour and physical capital, across different economic agents. Table 2.19 shows that although the state sector accounted for the large average share of 41.12 per cent per year in the total number of employees, its workforce was declining at the average rate of -2.24 per cent per year. The number of people working in the non-state and the foreign-invested sectors grew annually at high average rates of 21.02 per cent and 22.73 per cent, respectively, suggesting that these sectors had played an increasingly important role in creating jobs over this period. It is worth noting that the proportion of people working in foreign-invested enterprises was less than half of that of either the state sector or the non-state sector. In regard to the distribution of fixed assets (including long-term investments), on average, the non-state sector, while accounting for the average share of 41.70 per cent per year in the total number of employees, has an average share of just 17.65 per cent per year in the total value of fixed assets. On the contrary, although the state sector was a declining sector in terms of the number of enterprises and the number of employees, it still maintained the largest share in the total value of fixed assets of the economy.

Although the non-state sector played a significant role in job creation during this period, it appeared to be the least productive sector in terms of net turnover per employee and profit per employee. It was also much less capital-intensive, as illustrated by fixed assets per employee, than the other sectors. As noted in GSO (2011a, p.67), the workforce of the non-state sector was dominated by unskilled, and hence less productive, workers. With respect to the state sector and the foreign-invested sector, the former was not significantly less capital-intensive than the latter. However, in terms of profit per employee, foreign-invested enterprises were much more profitable than domestic enterprises. In comparison to the non-state sector and the foreign-invested sector, the state sector remained the most important taxation-source for the national budget.[9]

From the perspectives of macroeconomic stability and overall development, the efficiency of SOEs which still maintain a substantial contribution to the national income is a matter of great importance to the Vietnamese economy. According to the World Bank (2003, p.54), Vietnam could not have achieved economic progress in terms of economic growth without "generally increased efficiency in the state sector" despite the fact that "there are both profitable and loss-making SOEs". Recently, GSO (2011a, p.67) notes that the productivity as well as the technology-intensity within the SOE sector has not reached the high level. In addition, "wasteful uses of resources" and "losses of resources" have been found in a number of SOEs. Furthermore, some SOEs allocated large amounts of their capital to economic activities that were not their registered activities, causing harmful effects on the capital market and the markets of goods and services in which those investments involved.

Among the three economic sectors, the service sector was the most productive and the most capital-intensive sector. It is noteworthy that the capital intensity of the industry sector was less than half of that of the service sector and even slightly lower than that of the agriculture sector,[10] despite the fact that the industry sector accounted for the large average share of 69.44 per cent per year of total number of employees and its workforce grew annually at the average rate of 12.02 per cent. This might indicate the labour-intensive characteristic of the industry sector. The largest proportion of taxes payable to the national budget came from

[9] This could be influenced by differences in the industrial sectors where SOEs and other enterprises are specialised, and hence differences in the tax rates applied to them. For example, while the standard CIT rate is 25 per cent, the CIT rate applicable to activities of prospecting, exploring and exploiting oil and gas and other rare natural resources is between 32 per cent and 50 per cent, depending on each project or business establishment. With respect to the tax system in Vietnam, which is applicable for both domestic and foreign-invested enterprises, consists of the following main taxes: corporate income tax (CIT), import-export duties, value added tax (VAT), special sales tax (excise tax), and personal income tax.

[10] Note that these are the registered agriculture, industry and service sectors.

industry enterprises. Although the agriculture sector was a labour-intensive sector, it was a declining sector in terms of wage employment.

Table 2.19: Distribution of Enterprises by Workforce Size and Economic Sectors (2000-2007, Average Per Year)

Sector	Share in Total Employees (%)	Growth of Workforce (%)	Share in Total Fixed Assets (%)	Fixed Assets per Employee (Million VND)	Net Turnover per Employee (Million VND)	Profit per Employee (Million VND)	Share in Total Taxes Payable to the National Budget (%)
By Ownership							
SOEs	41.12	-2.24	52.56	228.61	365.32	19.62	46.11
NSEs	41.70	21.02	17.65	65.99	270.84	4.50	15.38
FIEs	17.18	22.73	29.79	262.39	384.93	52.29	38.51
By Economic Sector							
AGR	5.08	-0.74	3.85	113.65	70.38	9.96	0.93
IND	69.44	12.02	52.85	113.16	213.49	17.12	59.08
SER	25.48	11.45	43.30	269.17	657.96	22.11	39.98

Source: Calculation using Data from GSO (2010)
Note: SOEs, NSEs and FIEs denote state-owned enterprises, non-state enterprises and foreign-invested enterprises, respectively. AGR, IND and SER denote the agriculture sector, the industry domestic sector and the services sector, respectively.

Data on accumulated FDI from 1988 to 2010 can provide a further insight on the distribution of FDI by economic sector and, especially, by the country of origin. As presented in Table 2.20, the largest proportions of FDI projects and registered capital were in the manufacturing sector. In regard to the country of origin, FDI has essentially come from Asian countries. By 2010, within the group of non-Asian countries, the United States has been the largest foreign investor of Vietnam. According to Menon (2009), the country-of-origin composition of FDI in Vietnam is much diversified compared with those of many other developing economies, and this is attributed to the possibility that there exist a wider range of investment opportunities available to attracting FDI in Vietnam's relatively larger economy.

Table 2.20: Distribution of FDI (Accumulation as of 31/12/2010)

Distribution of FDI by Economic Sector			Distribution of FDI by Country of Origin	
Economic Sector	Share in Total Number of Projects (%)	Share in Total Registered Capital (%)	Top 20, by Total Number of Projects	Top 20, by Total Registered Capital
Agriculture, forestry and fishery	3.84	1.59	South Korea	Taiwan
Mining and quarrying	0.55	1.51	Taiwan	South Korea
Manufacturing	59.26	48.90	Japan	Singapore
Electricity, gas, stream and air conditioning supply	0.51	2.50	Singapore	Japan
Water supply, sewerage, waste management and remediation	0.19	0.03	China	Malaysia
Construction	5.67	5.96	Hong Kong	British Virgin Islands
Wholesale and retail trade; Repair of motor vehicles and motorcycles	4.15	0.85	United States	United States
Transport and storage	2.44	1.64	British Virgin Islands	Hong Kong
Accommodation and food services	2.42	5.85	Malaysia	Cayman Islands
Information and communication	5.26	2.48	France	Thailand
Financial, banking and insurance	0.60	0.68	Australia	Netherlands
Real estate	2.84	24.69	Thailand	Brunei
Professional, scientific and technical activities	7.95	0.36	Germany	Canada
Administrative and support services	0.79	0.09	Netherlands	China
Education and training	1.09	0.18	United Kingdom	France
Human health and social work	0.60	0.56	Brunei	Samoa
Arts, entertainment and recreation	0.99	1.79	Canada	United Kingdom
Others	0.84	0.33	Denmark	Cyprus
			Samoa	Switzerland
			Switzerland	Australia

Source: Calculation using Data from GSO Website
Note: The registered capital includes supplementary capital to the licensed projects.

2.3 DEVELOPMENT OF PHYSICAL INFRASTRUCTURE

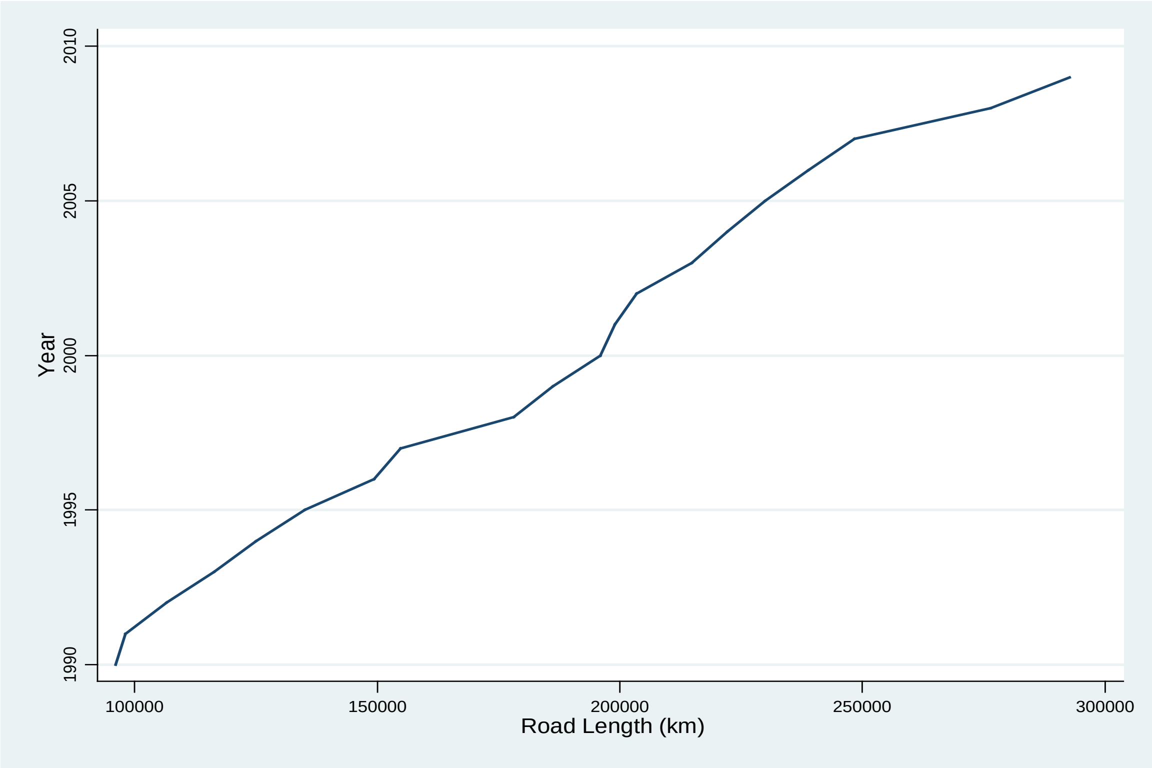
The World Bank (2000, p.1) notes that "the dramatic increase in the provision of infrastructure services in Vietnam since the late 1980s greatly facilitated rapid growth in exports and Gross Domestic Product (GDP)." Table 2.21 presents outlays of public spending over the period 2000-2007, showing a strong focus on infrastructure investment. Specifically, the proportion of infrastructure was highest in comparison to other categories of public spending. Public investment in infrastructure was close to 10 per cent of GDP in 2007. According to the World Bank (2009), this level of public infrastructure investment is very high by international standards.

Table 2.21: Structure of Public Outlays (2000-2007)

	2000	2001	2002	2003	2004	2005	2006	2007
	% of TOTAL SPENDING							
Capital Spending	27.19	31.00	30.51	32.91	30.87	30.15	28.68	28.08
Infrastructure	24.06	27.85	27.49	30.04	28.83	27.73	26.32	26.90
Others	3.13	3.16	3.02	2.87	2.04	2.42	2.36	1.18
Socio-Economic Spending	56.74	55.14	52.66	52.77	50.42	50.37	52.54	53.06
Education	11.63	11.89	12.04	12.63	11.83	10.89	12.12	13.46
Health	3.17	3.24	3.14	2.96	2.81	2.90	3.74	4.11
Pension	9.86	10.34	8.92	9.08	8.07	6.76	7.19	9.16
Others	32.08	29.66	28.55	28.10	27.71	29.83	29.49	26.33
Reserves	0.78	0.65	0.36	0.06	0.04	0.03	0.04	0.05
	% of GDP							
TOTAL SPENDING	24.67	26.96	27.66	29.54	29.94	31.30	31.62	34.92
Capital Spending	6.71	8.36	8.44	9.72	9.24	9.44	9.07	9.81
Infrastructure	5.93	7.51	7.60	8.87	8.63	8.68	8.32	9.39
Others	0.77	0.85	0.84	0.85	0.61	0.76	0.75	0.41
Socio-Economic Spending	14.00	14.87	14.57	15.59	15.10	15.77	16.61	18.53
Education	2.87	3.21	3.33	3.73	3.54	3.41	3.83	4.70
Health	0.78	0.87	0.87	0.88	0.84	0.91	1.18	1.44
Pension	2.43	2.79	2.47	2.68	2.42	2.11	2.27	3.20
Others	7.91	8.00	7.90	8.30	8.30	9.34	9.32	9.19

Source: Calculation using Data from GSO Website

Figure 2.1: Road Expansion



Source: Calculation using Data from Ministry of Transport

While a physical indicator, for example road density defined as total road length per 1,000 km2, performs as a proxy for transport infrastructure endowment derived from the supply aspect of infrastructure, the volumes of freight and passenger traffic can be an alternative proxy derived from the demand aspect of infrastructure. In regard to the physical indicator, Figure 2.1 illustrates that the road network has expanded in length more than two-folds since 1990.

Roads, together with maritime, were the main means of transport for passengers and freight, as can be seen in Table 2.22. There were some changes in the influence of the different means of transport, which can be illustrated by changes in the share of each mean of transport in total traffic volume, during the period 1995-2010. For road transport, the statistics indicate that its influence in freight traffic was not as great as that in passenger traffic. This, however, should be interpreted with caution. Geographically, Vietnam has a coastline of about 3,260 kilometres, excluding islands, and the system of seaports is located throughout the country from the North to the South. Therefore, under certain circumstances, such as transporting exports and imports, maritime could be a good alternative to roads as a mean of freight transport. However, since the statistics reported for each mean of transport do not differentiate between domestic and foreign destinations of transport, the aggregate data presented in Table 2.22 could not be very supportive to the assertion that a very large proportion of exports and imports had been transported by sea over the period observed.

With respect to the destination of freight traffic, as shown in Table 2.23, there was an increased proportion of foreign destination in total traffic volume, and the volume of freight traffic to foreign destination grew annually much faster than that to domestic destination. This should in part be the result of the opening-up policy which has encouraged the external trading activities. Table 2.23 also provides figures illustrating the influence of each economic agent in transport services. In particular, the state sector still maintained a dominant role in freight traffic.

Table 2.22: Freight Traffic and Passenger Traffic by Means of Transport (1995-2010)

Year	Freight Traffic by Mean of Transport (%)										Passenger Traffic by Mean of Transport (%)							
	Railways		Roads		Waterways		Maritime		Aviation		Railways		Roads		Waterways		Aviation	
	Share	Growth	Share	Growth	Share	Growth	Share	Growth	Share	Growth	Share	Growth	Share	Growth	Share	Growth	Share	Growth
1995	5.66	...	16.38	...	28.05	...	49.61	...	0.29	...	8.85	...	66.13	...	8.04	...	16.98	...
1996	4.35	-3.83	14.75	12.77	23.34	4.21	57.28	44.58	0.28	20.07	8.55	5.97	68.17	13.05	8.35	13.92	14.93	-3.57
1997	3.38	-8.93	13.69	8.62	22.94	14.99	59.72	22.04	0.27	12.23	8.92	9.54	68.67	5.82	8.29	4.36	14.12	-0.66
1998	2.95	-10.72	14.36	7.24	27.97	24.74	54.46	-6.73	0.25	-3.00	8.77	2.66	69.62	5.79	8.26	3.97	13.34	-1.40
1999	2.89	5.59	14.10	6.10	27.62	6.67	55.18	9.44	0.21	-9.52	8.92	7.07	69.75	5.44	8.08	2.95	13.25	4.53
2000	3.51	35.25	14.33	12.93	25.79	3.76	56.17	13.12	0.21	8.15	9.86	17.56	68.92	5.17	7.73	1.78	13.50	8.44
2001	3.25	5.08	14.54	15.24	26.81	18.06	55.14	11.47	0.25	38.65	9.62	7.07	65.67	4.55	7.56	7.29	17.15	39.42
2002	3.45	16.41	15.37	16.14	22.96	-5.91	57.98	15.56	0.25	8.60	9.39	7.91	65.05	9.41	7.51	9.84	18.05	16.21
2003	3.41	13.96	15.42	15.66	19.36	-2.79	61.56	22.39	0.26	22.58	9.17	10.06	68.63	18.99	6.17	-7.37	16.03	0.15
2004	3.03	0.73	16.51	21.08	18.14	5.96	62.06	14.02	0.26	11.97	8.55	7.55	66.97	12.50	6.17	15.28	18.31	31.71
2005	2.93	7.43	17.54	18.27	17.87	9.65	61.43	10.15	0.24	1.48	7.91	4.26	66.91	12.65	5.91	7.88	19.28	18.76
2006	3.04	16.86	18.09	16.24	16.60	4.69	62.05	13.87	0.24	12.58	6.78	-5.02	68.17	12.87	4.99	-6.39	20.05	15.21
2007	2.88	12.65	18.27	20.01	16.49	18.00	62.16	19.00	0.21	3.90	6.48	7.52	68.70	13.32	4.39	-1.19	20.43	14.55
2008	2.41	7.43	16.18	13.47	14.39	11.84	66.85	37.83	0.17	5.61	5.83	-2.13	69.35	9.82	4.15	3.01	20.66	10.02
2009	1.94	-7.35	15.87	12.94	15.70	25.66	66.33	14.27	0.16	7.10	4.86	-9.26	72.19	13.44	3.58	-6.10	19.37	2.20
2010	1.81	2.37	16.59	14.90	14.41	0.90	67.00	11.00	0.20	35.57	4.56	8.15	70.55	12.50	3.25	4.50	21.64	28.55
Average 95-10	3.18	6.20	15.75	14.11	21.15	9.36	59.69	16.80	0.23	11.73	7.94	5.26	68.34	10.36	6.40	3.58	17.32	12.27
Average 00-07	3.19	13.55	16.26	16.95	20.50	6.43	59.82	14.95	0.24	13.49	8.47	7.11	67.38	11.18	6.30	3.39	17.85	18.06

Source: Calculation using Data from GSO Website

Note: "Share" is the share of the relevant mean of transport in total volume of freight (or, passenger) traffic. "Growth" is the growth rate of the volume of freight (or, passenger) traffic of the relevant mean of transport.

Table 2.23: Freight Traffic and Passenger Traffic by Destination and by Economic Agent (1995-2010)

Year	Freight Traffic											Passenger Traffic						
		by Destination				by Economic Agent						Overall Growth	by Economic Agent					
	Overall Growth	Domestic		Foreign		State		Non-State		Foreign-Invested			State		Non-State		Foreign-Invested	
		Share	Growth	Share	Growth	Share	Growth	Share	Growth	Share	Growth		Share	Growth	Share	Growth	Share	Growth
1995	...	58.12	...	41.88	...	69.05	...	30.95	...	0.00	...	...	...	...	...	...	...	...
1996	25.23	50.32	8.43	49.68	48.55	73.30	32.94	26.70	8.04	0.00	...	32.94	...	...	...	...	...	...
1997	17.04	46.81	8.87	53.19	25.32	74.89	19.58	25.11	10.08	0.00	...	19.58	...	...	...	...	...	...
1998	2.27	57.64	25.93	42.36	-18.54	71.71	-2.07	28.29	15.22	0.00	...	-2.07	...	...	...	...	...	...
1999	8.02	61.90	16.01	38.10	-2.85	72.24	8.83	27.76	5.99	0.00	...	8.83	...	...	...	...	...	...
2000	11.14	58.39	4.85	41.61	21.36	73.58	13.19	26.42	5.79	0.00	...	13.19	34.15	...	65.85	...	0.00	...
2001	13.54	54.06	5.12	45.94	25.36	73.93	14.08	23.37	0.45	2.70	...	14.08	37.46	20.36	62.02	3.34	0.51	...
2002	9.90	54.36	10.51	45.64	9.19	74.59	10.88	22.81	7.24	2.61	6.24	10.88	38.90	14.72	60.63	7.98	0.47	0.77
2003	15.29	49.82	5.66	50.18	26.75	67.57	4.43	21.93	10.84	10.51	364.48	4.43	37.34	8.25	62.09	15.50	0.56	35.17
2004	13.09	48.05	9.07	51.95	17.08	66.70	11.64	22.11	14.03	11.19	20.47	11.64	39.67	22.49	59.76	10.96	0.57	16.62
2005	11.30	39.32	-8.93	60.68	30.00	63.60	6.13	30.35	52.78	6.05	-39.87	6.13	39.60	12.54	59.36	12.02	1.04	106.39
2006	12.73	31.39	-10.02	68.61	27.47	62.32	10.45	33.90	25.92	3.78	-29.46	10.45	38.15	6.72	61.00	13.82	0.85	-9.45
2007	18.79	32.00	21.11	68.00	17.72	65.34	24.54	31.18	9.26	3.48	9.38	24.54	38.08	12.24	61.04	12.53	0.88	15.99
2008	28.15	34.70	38.97	65.30	23.07	69.54	36.41	28.03	15.20	2.43	-10.70	36.41	37.10	5.99	61.82	10.16	1.08	34.30
2009	15.16	33.00	9.52	67.00	18.16	71.41	18.26	27.20	11.76	1.39	-34.17	18.26	35.47	4.18	62.67	10.49	1.86	87.05
2010	9.90	31.00	3.24	69.00	13.19	...	...	...	...	...	...	...	...	...	...	...	...	...
Average 95-10	14.10	46.30	9.89	53.70	18.79	69.98	14.95	27.07	13.76	2.94	35.79	14.95	NA	NA	NA	NA	NA	NA
Average 00-07	13.22	45.92	4.67	54.08	21.87	68.45	11.92	26.51	15.79	5.04	55.20	11.92	37.92	13.90	61.47	10.88	0.61	27.58

Source: Calculation using Data from GSO Website

Note: "Share" is the share of the relevant type of economic agent in total volume of freight (or, passenger) traffic. "Growth" is the growth rate of the volume of freight (or, passenger) traffic of the relevant type of economic agent.

Despite a large proportion of government expenditure that has been allocated to infrastructure, the infrastructure foundation of Vietnam remains extremely weak. According to the Global Competitiveness Reports, Vietnam's infrastructure quality has been given very low ratings. For example, the quality of Vietnamese infrastructure was ranked 90 in the 2006-2007 Report, 89 in the 2007-2008 Report, 97 in the 2008-2009 Report, and 111 in the 2009-2010 Report. As pointed out by the World Bank (2010, p.326), infrastructure remains "a bottleneck of the economy, impeding import, export, and circulation of commodities in Vietnam." In-migration from the rural to urban areas, which according to GSO (2011b) has been experiencing an upward trend, is another factor that puts further pressure on the congested and underdeveloped urban infrastructure system.

The World Bank (2000) highlights two sources of Vietnam's infrastructure problems, namely, a lack of infrastructure assets and the poor performance of SOEs supplying infrastructure services such as transport services and energy and water supply services. Among reasons for the poor performance of these SOEs, the most important are price controls and a lack of competition from private enterprises in the infrastructure sector (World Bank, 2000). According to Chung *et al.* (2010), in recent years, there have been several legal documents which aimed at creating the legal environment and thereby encouraging the private sector participation in large-scale infrastructure projects such as seaports and international airports. The private sector participation is expected to provide a large volume of additional financial resources for infrastructure investment but also to improve efficiency in the supply of infrastructure services (Chung *et al.*, 2010).

Another type of infrastructure examined in the empirical research is higher-education infrastructure. In regard to the Vietnamese education system, apart from pre-primary education, vocational training and postgraduate education, there are four education levels:

primary education, which requires 5 years to complete, secondary education, including lower and upper, which requires a total of 7 years to complete, and undergraduate education, of which the number of academic years to complete is subject-specific. As shown in Table 2.21, the ratio of public spending on education to GDP increased from 2.87 per cent in 2000 to 4.70 per cent in 2007. In addition, education accounted for the second-largest proportion in total public spending, which was just smaller than that of infrastructure. The composition of public spending generally suggests that the policy makers could have placed a great importance on infrastructure and education.

Table 2.24 shows that growth in the number of education institutions can be seen at all educational levels, although the average growth rate of colleges and universities was much higher than those of primary and secondary schools. The number of education institutions is also divided by the number of inhabitants to consider whether the provision of education infrastructure met the schooling demand. In some years during the period 1995-2010, we can also see negative per capita growth rates of education institutions at the primary and secondary levels. To some extent, these figures could indicate that, at some specific educational levels, the provision of education infrastructure lagged behind the growth of schooling demand proxied by the overall growth of population. On the contrary, the average growth rate of the number of colleges and universities was slightly higher than the average growth rate of GDP – a conventional indicator of economic development – over the same period. The figures reported in Table 2.24 are obviously unable to capture the quality of the education infrastructure as well as the capacity of the higher-education institutions in providing qualified labour. However, according to GSO (2011a), the training activities at vocational schools, colleges and universities have been largely driven by the capacity of those institutions in providing training services, not by the demand of the labour market.

Consequently, their graduates face difficulties in seeking jobs while enterprises face difficulties in recruiting appropriate candidates as employees.

Table 2.24: Growth of Education Infrastructure (1995-2010)

Academic Year	Growth in the Number of Education Institutions			Growth in the Number of Education Institutions per 1,000 Inhabitants		
	Primary Schools	Secondary Schools	Colleges and Universities	Primary Schools	Secondary Schools	Colleges and Universities
1995-1996	...	...	...	...	...	...
1996-1997	3.79	2.79	-11.93	2.15	1.16	-13.32
1997-1998	5.10	3.03	14.58	3.47	1.43	12.81
1998-1999	3.88	2.61	11.82	2.30	1.04	10.11
1999-2000	1.95	3.32	6.50	0.43	1.78	4.92
2000-2001	2.53	3.22	35.88	1.16	1.84	34.06
2001-2002	0.32	4.51	7.30	-0.94	3.20	5.96
2002-2003	1.87	3.00	5.76	0.70	1.82	4.54
2003-2004	1.29	2.95	5.94	0.12	1.76	4.72
2004-2005	1.20	2.44	7.48	-0.01	1.22	6.20
2005-2006	1.17	1.95	20.43	0.00	0.77	19.04
2006-2007	0.99	1.75	16.25	-0.12	0.63	14.96
2007-2008	0.67	1.61	14.60	-0.42	0.52	13.36
2008-2009	0.79	0.76	6.50	-0.28	-0.31	5.38
2009-2010	0.80	1.32	2.54	-0.26	0.26	1.46
2010-2011	0.46	0.87	2.73	-0.58	-0.18	1.66
Average 1995-2010	1.79	2.41	9.76	0.51	1.13	8.39
Average 2000-2007	1.26	2.68	14.20	0.06	1.47	12.85

Source: Calculation using Data from GSO Website

Since the 1990s, industrial zones (IZs) and export processing zones (EPZs) have been established to stimulate industrial development.[11] According to Decree No. 192-CP dated 28/12/1994 of the Government of Vietnam, economic activities taking place in the IZs include construction and supply of infrastructure, manufacturing and assembly of industrial products which are to be exported and/ or distributed in the domestic market, and other

[11] As defined in the Foreign Investment Law and, later, in the Investment Law, an Export Processing Zone (EPZ) is an industrial zone which has definite geographical boundaries and specialise in the production of exports and the provision of services for the production of exports and export activities. An Industrial Zone (IZ) is a zone which has definite geographical boundaries and specialises in the production of industrial goods and the provision of services for industrial production. As noted previously, these zones are also classified as the geographical areas entitled to investment preferences.

services supporting industrial activities. Accordingly, all enterprises, regardless of their ownership, can be located in IZs subject to their budget constraints and their specialised industries, as well as subject to the availability of lots being ready for enterprises to move in.

By 2007, Vietnam had 202 IZs which, as shown in Table 2.25, are essentially concentrated in three regions, namely, the Red River Delta, the Central Coast and the Southeast. In particular, these three regions accounted for 77.72 per cent of the total number of IZs. Also by 2007, there were 6 EPZs, which are located in the Red River Delta and the Southeast. The uneven geographical distribution of these zones indeed could not be attributed to any biased public policies. More precisely, the open-up policies, including the IZ and EPZ policies, have been applied to all regions and provinces at the same time, and the establishment of the zones was essentially subject to economic conditions of the area where they are located.

Table 2.25: Regional Distribution of IZs and EPZs (2007, Number of Zones)

	Red River Delta	Northern Mountain	Central Coast	Central Highlands	Southeast	Mekong River Delta
IZs	45	10	49	7	63	28
EPZs	2	0	0	0	4	0

Source: GSO (2011a)

2.4 SPATIAL STRUCTURE OF THE ECONOMY

The following discussion is based on the average per year statistical indicators for the period 2000-2007 presented in Tables 2.26 and 2.27. As noted in the previous chapter, Vietnamese provinces can be grouped into six geographical regions, namely, the Red River Delta (henceforth, RRD), the Northern Mountain (henceforth, NM), the Central Coast (henceforth, CC), the Central Highlands (henceforth, CH), Southeast (henceforth, SE) and the Mekong River Delta (henceforth, MRD).

RRD was the most populous region, accounting for 23.07 per cent of the national population. In addition, with the population density that is slightly above 4 times the national average, RRD appeared to be the densest region of the country. Nevertheless, RRD experienced a net migration rate of -1.51, indicating that the number of in-migrants is smaller than that of out-migrants. In regard to geographical concentration of economic activities, RRD was the region having the second-largest share of the national number of enterprises. The region was also the second-largest contributor to national GDP as well as the second-largest owner of production inputs, such as physical capital and labour. With respect to economic structure, the proportion of agriculture in GDP was slightly higher than the national average, while the measure of trade openness, i.e. the ratio of exports to GDP, was below. Regarding the role of the state and the private sectors in this regional economy, the proportion of private sector employment in total employment was slightly below the national average, whereas the share of SOEs in physical capital stock was slightly above. Urbanisation rate was below the national average, and this could be explained by the fact that agriculture still maintained an important role in the regional economic structure. In contrast, the region's road density was about 2 times the national average. With respect to higher-education infrastructure, 39.70 per cent of the national number of higher-education institutions was located in this region.

In comparison to other RRD provinces, Ha Noi, a special metropolis located in the region, had the highest share of 3.73 per cent in the national population, and its population density was slightly above 13 times the national average. Ha Noi also had an extremely high rate of 14.36 of net migration while all other provinces belonging to the same region had negative net migration rates. With respect to the geographical concentration of economic activities, the number of enterprises located in Ha Noi was even higher than the number of enterprises located in all NM provinces, or the number of enterprises located in all CC provinces, or the number of enterprises located in all CH provinces. Similarly, the shares of Ha Noi in the national stock of physical capital and number of employees were higher than those of NM, or CC, or CH. Compared with the private sector, the state sector had a larger contribution to total employment as well as possessed a larger proportion of physical capital stock. While the proportion of private sector employment in total employment was below the national average, the share of SOEs in physical capital stock was above. At the province level, Ha Noi was the third-largest contributor to the national GDP. As shown by economic structure indicators, Ha Noi was a non-agriculture-based provincial economy which also had a high ratio of exports to GDP. Ha Noi also had a very high rate of urbanisation compared with the national average. Regarding infrastructure, Ha Noi contributed significantly to the large share of RRD in the national number of higher-education institutions. Specifically, Ha Noi accounted for 22.76 per cent of the relevant indicator.

Although CC accounted for the second-largest share of 22.84 per cent of the national population, its population density was below the national average. A majority of provinces belonging to CC had negative net migration rates. The performance of other economic indicators was not as impressive as that of RRD and SE, but still better than that of others regions, notably, NM and CH. With respect to infrastructure provision, the density of roads installed in this region was below the national average, which was just higher than that of NM and CH; whereas, the share of the region in the number of higher-education institutions was just smaller than those of RRD and SE.

The 14.63 per cent share of SE in the national population was much smaller than that of RRD, CC or MRD. However, SE is the second-densest region with population density slightly above 3 times the national average. SE also had a high rate of net migration. In addition, SE was the most dynamic region since a substantial proportion of enterprises, and hence their physical capital and labour, was located in this region. As a result, SE was the largest regional contributor to national GDP and exports. Compared with the national average, SE had a more non-agriculture-based economic structure and was more open to external trade. Also, in comparison to the national average, the private sector played an influential role in employment in the SE while the share of SOEs in physical capital stock was not as significant as that could be seen in other regions. Finally, infrastructure indicators, i.e. the density of roads and the number of higher-education infrastructure, were higher than the national average.

Ho Chi Minh City (henceforth, HCMC or the City) is another special metropolis, which is located in SE. Compared with other SE provinces, HCMC had the highest share of 7.34 per cent in the national population, and its population density was more than 11-times higher than the national average level. Also, HCMC had an extremely high rate of 14.91 of net migration. The population density and the net migration rate of HCMC were relatively similar to those of Ha Noi. The road network installed in HCMC was also as dense as that in Ha Noi,[12]

[12] To some extent, a high network density does not necessarily indicate a sufficient provision of infrastructure. During the period observed, while Ha Noi and Ho Chi Minh City had a high density of roads, the level of roads distributed to their inhabitants measured by road length per capita was significantly low compared with other provinces. The level of Ho Chi Minh City was lower than that of Ha Noi.

although the rate of urbanisation of the former was higher than that of the latter. Specifically, the urbanisation rate of HCMC was slightly above 3 times the national average, while the rate of Ha Noi was slightly above 2 times. This could be explained partially by the share of agriculture in GDP of these two special metropolises. At the province level, HCMC was the largest contributor to national GDP and exports. Specifically, on average, the City contributed 18.43 per cent to national GDP and 42.82 per cent to national exports. The City also maintained a high ratio of exports to GDP, which was slightly above twice the national average. With respect to the geographical concentration of economic activities, HCMC was the most dynamic province of the country, indicated by 25.12 per cent of enterprises located in the City as well as the rate of 27.41 per cent of physical capital stock possessed, and the rate of 24.89 per cent of the workforce employed, by the City's enterprises. On the contrary to Ha Noi, and compared with the relevant national-average indicators, the role of the City's private sector in job creation was more influential than that of the state sector. In addition, the proportion of physical capital stock owned by the private sector was larger than that of the state sector. In regard to the number of higher-education institutions, Ha Noi accounted for a share that was larger than that of HCMC. Provided that the number of higher-education institutions per 1,000 inhabitants could be considered an indicator of whether the provision of higher-education infrastructure is likely to meet the schooling demand, then the relevant indicators for Ha Noi and HCMC, being much lower than the national average, suggested a congestion in using infrastructure in these special metropolises; though, the problem seemed to be more serious in HCMC. Generally, if HCMC performed more impressively than Ha Noi in a wide range of economic indicators, the provision of roads and higher-education infrastructure in the former was relatively less impressive than that in the latter. In this sense, although the infrastructure provision in these two special metropolises seem to be lagged behind the growth of infrastructure demand, the problem of congestion was likely more serious in HCMC. This could be partially explained by the concentration of economic activities in the City which was denser than that in Ha Noi.

While MRD has the population share that is nearly double than that of SE, the former is less dense than the latter. Nevertheless, its population density was still above the national average. This is contrary to CC which accounted for more than 20 per cent of the national population, as in the case of MRD, but had a below national-average level of population density. The geographical concentration of economic activities in MRD was not as high as in other regions, notably RRD and SE. However, the region still had a higher share in the national enterprises, in comparison to NM, CC, and CH. In MRD, the private sector had an important role in employment and, at the same time, SOEs still maintained a larger share of physical capital stock. This pattern is totally different with those of other regions where the proportion of jobs created by private enterprises was below the national average, while the proportion of physical capital stock possessed by SOEs was above. With respect to economic structure, MRD was an agriculture-based regional economy. Yet, the shares of MRD in national GDP and exports were much higher than those of the regions where agriculture still accounted for an influential proportion of GDP. This could be explained by the pattern of the agriculture-based economy of MRD which was export-led growth. In regard to infrastructure, the density of the road network of MRD was just slightly-above the national average. The rate of urbanisation in this region is below the national average, and this is understandable for the agriculture-based region.

A comparison between regions in regard to a wide range of indicators presented in Tables 2.26 and 2.27 reveals that NM and CH were the most backward regions. For NM, only 13.15 per cent of the national population lived in this region and the population density of the region was below the national average. All provinces belonging to NM experienced negative net migration rates. During the period observed, the least populous region was CH which accounted for only 5.69 per cent of the national population. It was the sparsest region, as indicated by the density of population which was not only below the national average but also lower than any other regions. However, CH experienced a positive net migration rate of 1.29. Economic indicators calculated for these two regions were also less impressive than those of other regions. With respect to infrastructure, the density of roads installed in these backward regions was significantly lower than the national average, and their shares in the national number of higher-education institutions were also very small in comparison to those of other regions. One difference between NM and CH is that, compared with the national average, the urbanisation rate was significantly lower for the former, but higher for the latter.

Main characteristics of the spatial economy of Vietnam of the period 2000-2007 can be summarised, at the region level, as follows. Population was unevenly distributed across regions and heavily concentrated in RRD, CC and MRD. In regard to population density, RRD was the densest region, followed by SE and MRD. SE and CH were the two regions experiencing positive net migration rates, although that positive rate of the former was much higher than that of the latter. With respect to the geographical concentration of economic activities, especially private-sector economic activities, SE was the most dynamic region, followed by RRD. A large proportion of industrial production was concentrated in SE and RRD as well. Urbanisation provided a mixed picture, which was somewhat associated with the economic structure of a specific region. The economic development level has also been very different between regions, and the different level of economic development can be seen from the disparities in regard to the contribution of each region to national GDP. Accordingly, SE maintained the role of the leading region in the Vietnamese economy, while the poorest region was NM. The two southern regions, i.e. SE and MRD, together comprised a substantially large share of national exports, indicating that the export activities of the nation had been concentrated in these regions. Finally, infrastructure and other resources, such as labour and physical capital, had been distributed unevenly throughout the country.

Table 2.26: Agglomeration and Spatial Distribution of GDP, Capital and Labour (2000-2007, Average Per Year)

	Share in National Population	Net Migration Rate	Share in National Number of Enterprises	Share in National Number of Non-State Enterprises	Share in National Number of Foreign-Invested Enterprises	Share in National Exports	Share in National GDP	Share in National Physical Capital Stock	Share in National Number of Employees
Red River Delta	**23.07**	**-1.51**	**26.86**	**26.39**	**21.78**	**7.08**	**22.33**	**29.45**	**29.72**
Ha Noi	3.73	14.36	15.03	14.90	12.80	2.09	7.01	17.82	14.16
Ha Tay	3.22	-3.12	1.40	1.35	0.77	0.31	1.92	0.78	1.30
Vinh Phuc	1.41	-2.97	0.70	0.68	0.94	0.39	1.22	0.76	0.72
Bac Ninh	1.21	-1.63	0.95	0.98	0.42	0.27	0.95	0.74	0.84
Quang Ninh	1.33	-0.42	1.19	1.12	0.76	0.31	1.48	1.59	2.76
Hai Duong	2.07	-3.91	1.23	1.22	1.20	0.43	1.74	1.08	1.40
Hai Phong	2.16	-0.76	2.68	2.53	3.80	2.01	2.87	3.87	4.16
Hung Yen	1.36	-1.41	0.59	0.57	0.66	0.51	1.08	0.55	0.87
Thai Binh	2.21	-2.91	0.90	0.88	0.16	0.27	1.41	0.43	1.03
Ha Nam	0.98	-5.79	0.46	0.45	0.08	0.14	0.61	0.57	0.42
Nam Dinh	2.30	-4.72	1.06	1.05	0.12	0.30	1.41	0.76	1.37
Ninh Binh	1.10	-4.89	0.67	0.67	0.06	0.07	0.63	0.49	0.69
Northern Mountain	**13.15**	**-1.24**	**5.81**	**5.47**	**2.10**	**1.72**	**6.53**	**3.17**	**5.86**
Ha Giang	0.82	-0.54	0.27	0.26	0.00	0.01	0.29	0.09	0.31
Cao Bang	0.62	-1.95	0.28	0.25	0.02	0.04	0.36	0.13	0.30
Bac Kan	0.35	-0.06	0.22	0.21	0.02	0.01	0.15	0.04	0.12
Tuyen Quang	0.87	-1.18	0.31	0.29	0.00	0.00	0.46	0.09	0.29
Lao Cai	0.75	-0.43	0.54	0.54	0.23	0.07	0.33	0.23	0.46
Yen Bai	0.88	-1.44	0.38	0.35	0.13	0.03	0.45	0.15	0.41
Thai Nguyen	1.34	-0.71	0.73	0.72	0.25	0.12	0.80	0.60	0.93
Lang Son	0.89	-3.02	0.41	0.38	0.19	0.80	0.59	0.18	0.26
Bac Giang	1.89	-3.58	0.68	0.65	0.36	0.19	0.84	0.24	0.55
Phu Tho	1.59	-1.63	0.94	0.90	0.63	0.37	0.95	0.91	1.27
Dien Bien & Lai Chau	0.94	-0.17	0.35	0.30	0.05	0.01	0.36	0.11	0.29

Son La	1.21	-0.33	0.30	0.25	0.05	0.01	0.43	0.23	0.31
Hoa Binh	1.01	-1.10	0.42	0.38	0.18	0.05	0.50	0.16	0.35
Central Coast	**22.84**	**-1.93**	**14.94**	**14.97**	**4.62**	**4.85**	**14.65**	**9.77**	**13.23**
Thanh Hoa	4.26	-4.63	1.39	1.34	0.21	0.26	2.53	1.73	1.58
Nghe An	3.57	-1.41	1.58	1.51	0.17	0.24	2.14	1.40	1.36
Ha Tinh	1.55	-2.41	0.65	0.63	0.07	0.12	0.87	0.25	0.45
Quang Binh	1.01	-0.94	0.78	0.80	0.04	0.09	0.47	0.41	0.54
Quang Tri	0.72	-3.20	0.55	0.53	0.07	0.08	0.39	0.28	0.37
Thua Thien Hue	1.32	-2.61	1.10	1.10	0.34	0.18	0.74	0.63	0.79
Da Nang	0.96	2.43	2.30	2.33	1.14	1.06	1.23	1.45	2.38
Quang Nam	1.73	-3.76	0.77	0.74	0.39	0.27	1.03	0.44	0.79
Quang Ngai	1.49	-4.80	0.68	0.68	0.07	0.08	0.78	0.43	0.50
Binh Dinh	1.82	-3.50	1.16	1.19	0.23	0.65	1.19	0.53	1.55
Phu Yen	1.02	-0.23	0.63	0.64	0.18	0.18	0.54	0.30	0.52
Khanh Hoa	1.35	-0.78	1.74	1.78	1.14	1.24	1.54	1.37	1.55
Ninh Thuan	0.67	0.87	0.36	0.36	0.10	0.08	0.41	0.15	0.24
Binh Thuan	1.37	-2.09	1.25	1.35	0.49	0.31	0.80	0.41	0.60
Central Highlands	**5.69**	**1.29**	**3.39**	**3.21**	**1.69**	**1.92**	**4.15**	**2.34**	**3.10**
Kon Tum	0.45	0.95	0.30	0.26	0.01	0.04	0.27	0.18	0.31
Gia Lai	1.40	1.83	0.74	0.70	0.08	0.22	0.72	0.88	0.93
DakLak & DakNong	2.49	1.30	1.18	1.11	0.12	1.30	1.95	0.71	1.26
Lam Dong	1.35	1.06	1.17	1.14	1.48	0.35	1.20	0.57	0.60
Southeast	**14.63**	**6.65**	**33.28**	**33.00**	**66.99**	**82.13**	**33.42**	**51.05**	**40.80**
Binh Phuoc	0.94	1.45	0.56	0.59	0.14	0.53	0.54	0.37	0.75
Tay Ninh	1.26	-2.24	0.86	0.84	1.53	1.08	1.32	0.87	0.84
Binh Duong	1.26	24.03	2.69	2.19	18.90	7.11	1.59	5.46	6.80
Dong Nai	2.72	-0.98	2.77	2.49	11.80	9.90	3.85	7.41	6.05
Ba Ria - Vung Tau	1.12	2.74	1.27	1.26	1.66	20.69	7.68	9.53	1.47
Ho Chi Minh City	7.34	14.91	25.12	25.64	32.96	42.82	18.43	27.41	24.89
Mekong River Delta	**20.61**	**-2.41**	**15.73**	**16.96**	**2.81**	**9.43**	**18.92**	**4.22**	**7.29**
Long An	1.70	-2.57	1.31	1.35	1.47	1.19	1.57	0.76	1.26
Dong Thap	2.00	-3.38	1.19	1.30	0.04	0.55	1.55	0.22	0.40

An Giang	2.59	-2.56	1.49	1.62	0.13	0.96	2.19	0.44	0.67
Tien Giang	2.02	-2.56	1.88	2.05	0.12	0.50	1.73	0.38	0.82
Vinh Long	1.26	-3.03	1.03	1.12	0.13	0.38	0.98	0.22	0.46
Ben Tre	1.58	-4.88	1.38	1.52	0.12	0.28	1.32	0.23	0.40
Kien Giang	1.96	-0.40	2.22	2.45	0.08	0.55	2.21	0.47	0.73
Can Tho & Hau Giang	2.33	-2.83	1.80	1.88	0.55	1.57	2.44	0.71	1.17
Tra Vinh	1.22	-1.61	0.58	0.62	0.02	0.17	0.95	0.13	0.23
Soc Trang	1.53	-2.17	0.89	0.97	0.02	1.15	1.39	0.25	0.43
Bac Lieu	0.98	-1.30	0.71	0.76	0.08	0.44	0.98	0.14	0.20
Ca Mau	1.44	-1.56	1.24	1.33	0.05	1.67	1.61	0.26	0.51

Source: Calculation using Data from GSO, Ministry of Education and Training, and Ministry of Transport

Note: Monetary values are deflated by provincial GDP deflator. Indicators of a region are the averages of the relevant indicators calculated for provinces belonging to that region. Net migration rate is the difference of immigrants and emigrants of an area in a period of time, divided by 1,000 inhabitants.

Table 2.27: Spatial Differentials in Economic Structure, Exports, Private Sector Development and Infrastructure Provision (2000-2007, Average Per Year)

	Share of Agriculture in GDP/ National Average	Employment by Private Sector/ National Average	Share of SOEs in Physical Capital Stock/ National Average	Population Density/ National Average	Urbanisation/ National Average	Road Density/ National Average	Higher-Education Institutions per 1,000 Inhabitants/ National Average	Share in National Number of Higher-Education Institutions
Red River Delta	**1.07**	**0.94**	**1.13**	**4.73**	**0.81**	**2.07**	**1.00**	**39.70**
Ha Noi	0.11	0.67	1.56	13.39	2.38	2.74	0.66	22.76
Ha Tay	1.31	0.93	0.73	4.86	0.36	2.06	0.90	3.61
Vinh Phuc	0.96	1.10	0.38	3.40	0.50	2.08	1.03	0.94
Bac Ninh	1.07	1.06	0.57	4.87	0.55	2.34	0.98	1.21
Quang Ninh	0.32	0.36	1.73	0.72	1.83	0.61	1.06	1.21
Hai Duong	0.99	0.97	0.92	4.15	0.57	1.80	1.10	1.47
Hai Phong	0.49	0.92	1.09	4.70	1.50	2.50	0.84	2.61
Hung Yen	1.44	1.07	0.55	4.88	0.40	2.12	1.03	1.62
Thai Binh	1.98	1.13	0.89	4.73	0.27	2.36	0.98	1.17
Ha Nam	1.35	1.08	2.03	3.77	0.29	2.11	1.27	0.93
Nam Dinh	1.37	1.03	1.58	4.61	0.56	1.60	0.97	1.44
Ninh Binh	1.48	0.92	1.53	2.62	0.57	2.47	1.11	0.73
Northern Mountain	**1.63**	**0.88**	**1.45**	**0.62**	**0.58**	**0.49**	**1.57**	**9.02**
Ha Giang	1.80	1.29	0.66	0.34	0.43	1.01	1.42	0.39
Cao Bang	1.52	1.01	1.36	0.31	0.55	0.48	2.11	0.39
Bac Kan	2.01	0.95	1.22	0.24	0.58	0.22	2.28	0.34
Tuyen Quang	1.59	0.72	1.51	0.49	0.41	0.62	1.42	0.48
Lao Cai	1.50	0.98	1.34	0.39	0.72	0.27	1.99	0.39
Yen Bai	1.51	0.93	1.41	0.42	0.75	0.49	1.59	0.48
Thai Nguyen	1.19	0.65	1.86	1.25	0.90	0.74	1.22	2.15
Lang Son	1.67	0.88	1.69	0.35	0.72	0.33	1.63	0.51
Bac Giang	1.85	0.78	1.62	1.63	0.34	0.62	1.06	0.95
Phu Tho	1.08	0.80	1.40	1.49	0.57	0.47	1.42	1.11

Dien Bien & Lai Chau	1.59	0.95	1.29	0.17	0.54	0.21	1.29	0.39
Son La	1.99	0.73	1.94	0.28	0.47	0.53	1.29	0.87
Hoa Binh	1.84	0.76	1.56	0.72	0.57	0.42	1.71	0.56
Central Coast	**1.28**	**0.86**	**1.55**	**0.89**	**0.99**	**0.63**	**1.11**	**15.53**
Thanh Hoa	1.26	0.79	0.88	1.27	0.37	0.60	1.31	0.79
Nghe An	1.44	0.64	1.68	0.72	0.43	0.29	1.28	1.96
Ha Tinh	1.65	0.75	2.01	0.85	0.44	0.63	1.36	0.51
Quang Binh	1.17	0.75	1.88	0.42	0.50	0.21	1.59	0.48
Quang Tri	1.52	0.90	2.10	0.50	0.97	0.46	1.52	0.48
Thua Thien Hue	0.78	0.74	1.30	0.86	1.23	0.77	1.03	1.82
Da Nang	0.24	0.75	1.85	2.52	3.17	2.60	0.66	3.32
Quang Nam	1.29	0.93	1.10	0.55	0.63	0.41	1.07	0.56
Quang Ngai	1.50	0.86	2.09	0.96	0.51	0.52	1.03	1.35
Binh Dinh	1.67	1.20	1.29	1.00	0.98	0.25	0.84	1.03
Phu Yen	1.21	0.81	1.63	0.67	0.77	0.44	1.01	0.88
Khanh Hoa	0.87	1.10	1.06	0.86	1.45	0.66	0.80	1.57
Ninh Thuan	1.85	0.71	1.69	0.66	1.10	0.55	1.04	0.39
Binh Thuan	1.51	1.16	1.08	0.58	1.31	0.40	1.00	0.39
Central Highlands	**2.30**	**0.74**	**1.87**	**0.33**	**1.14**	**0.29**	**1.16**	**3.16**
Kon Tum	1.95	0.72	2.09	0.16	1.26	0.22	1.45	0.56
Gia Lai	2.05	0.66	2.19	0.30	1.02	0.38	1.02	0.48
DakLak & DakNong	2.73	0.52	1.96	0.42	0.79	0.22	1.06	1.03
Lam Dong	2.47	1.06	1.26	0.46	1.47	0.35	1.12	1.09
Southeast	**0.87**	**1.08**	**0.89**	**3.00**	**1.41**	**1.42**	**0.77**	**23.27**
Binh Phuoc	2.26	0.59	2.14	0.45	0.62	0.61	0.91	0.26
Tay Ninh	1.58	0.97	1.31	1.04	0.56	0.82	1.26	0.48
Binh Duong	0.47	1.40	0.34	1.54	1.16	1.78	0.60	1.15
Dong Nai	0.68	1.33	0.34	1.53	1.21	1.32	0.68	1.55
Ba Ria - Vung Tau	0.16	0.96	0.48	1.86	1.73	1.23	0.76	0.87
Ho Chi Minh City	0.07	1.20	0.75	11.58	3.16	2.74	0.42	18.97
Mekong River Delta	**1.94**	**1.10**	**1.26**	**1.87**	**0.72**	**1.03**	**0.88**	**9.32**
Long An	1.73	1.44	0.25	1.25	0.64	0.99	0.85	0.51

Dong Thap	2.29	1.06	1.61	1.96	0.60	1.41	0.84	0.87
An Giang	1.40	1.04	1.19	2.43	0.94	1.03	0.84	0.48
Tien Giang	1.83	1.26	0.97	2.70	0.51	1.13	0.72	0.79
Vinh Long	1.88	1.25	1.19	2.82	0.57	1.18	1.11	2.08
Ben Tre	2.29	1.06	1.70	2.22	0.35	1.41	0.86	0.53
Kien Giang	1.97	1.14	1.26	1.02	0.92	0.65	0.81	0.81
Can Tho & Hau Giang	1.10	0.92	1.44	2.57	1.20	1.50	0.76	1.55
Tra Vinh	2.52	0.98	1.93	1.76	0.53	0.57	0.98	0.65
Soc Trang	2.35	1.18	0.90	1.53	0.71	0.87	0.92	0.36
Bac Lieu	2.07	1.00	1.39	1.26	0.97	0.87	0.89	0.43
Ca Mau	1.87	0.88	1.25	0.90	0.75	0.73	0.97	0.26

Source: Calculation using Data from GSO, Ministry of Education and Training, and Ministry of Transport

Note: Monetary values are deflated by provincial GDP deflator. Indicators of a region are the averages of the relevant indicators calculated for provinces belonging to that region. "Employment by Private Sector" is the proportion of people working in private sector enterprises in total number of employees. "Urbanisation" is proxied by the proportion of people living in urban areas. "/ National Average" indicates that the provincial indicator is proportionate to the relevant indicator calculated at the national level.

2.5 SUMMARY

By the end of 1985, the post-wartime economy of Vietnam completely fell into crisis. Since 1986, reforms have led to a number of important economic achievements. Before 2008, the Vietnamese economy grew annually at an average rate of around 7 per cent. Also, the economy had witnessed a fundamental transformation in its GDP structure towards industrialisation, an impressive growth and diversification of external economic relations, and an explosion of private entrepreneurship. With respect to the structural transformation in terms of ownership, the state sector still accounted for a very large share in total GDP, indicating the influential role of SOEs in the economy. The development of the private sector was most impressive in terms of the increased number of non-SOEs, i.e. non-state enterprises and foreign-invested enterprises, and hence the increased proportion of employment by those enterprises. Yet, the proportion of large-scale non-SOEs was extremely small in comparison to the substantial proportion of micro and small non-SOEs. The labour market was characterised in part by a large labour pool and a substantially large proportion of unskilled and semi-skilled labour. Regarding infrastructure, despite the great emphasis of government expenditure on infrastructure investment, the infrastructure foundation remained extremely weak especially in terms of quality. Finally, the regional development gap remains an issue to be solved, and the economic literature says that investment in infrastructure is a useful policy instrument for narrowing the development gap across sub-national areas. However, over the period 2000-2007, resources for economic development, such as infrastructure, labour and physical capital, had not been distributed evenly throughout the country.

CHAPTER 3

INFRASTRUCTURE AND ECONOMIC GROWTH

3.1 GENERAL BACKGROUND

Infrastructure is one of the most essential prerequisites for economic development, which is a conventional proposition in the economic literature and also provides a justification for public infrastructure investment. The most direct approach to analysing the contribution of infrastructure to economic development would be to examine its growth impact. Theoretically, exogenous and endogenous approaches of growth modelling capture differently the growth impact of public capital. In the exogenous growth approach, pioneered by Solow (1956), technical progress is the driver of long-run growth, and an increased stock of public capital can only have a "transitory" impact on growth. The growth impact of public capital was first endogenised in the production function by Barro (1990). In comparison to the Solow model, which explains the long-run growth by exogenous factors, the Barro model yields constant returns to scale and, consequently, there exists a positive long-run growth rate that can be explained endogenously. The growth model of Barro (1990) has been extended in the ensuing literature of endogenous growth models to incorporate various aspects which could affect the growth impact of public capital.

While the theoretical link between infrastructure and economic growth has been well established in the endogenous growth literature, empirical studies have provided conflicting results on the growth impact of infrastructure. Empirical evidence varies dramatically across studies, ranging from no growth impact to a significantly positive impact of infrastructure on economic growth. Sources of the conflicting empirical results could include variations across studies in regards to empirical model specification, econometric techniques, measurement of infrastructure, time and geographical dimension of data, and, especially for analyses at the sub-national level, attempts to control for the potential existence of cross-region spillovers from infrastructure.

As noted by the World Bank (2000, p.1), "the dramatic increase in the provision of infrastructure services in Vietnam since the late 1980s greatly facilitated rapid growth in exports and Gross Domestic Product (GDP)." Chapter 2 shows that, during the period 2000-2007, a high annual growth rate of GDP had been observed at the same time as a strong focus of public spending on infrastructure. In particular, the Vietnamese economy achieved an average annual real GDP growth rate of 7.63 per cent between 2000 and 2007. The outlays of public spending indicate that the proportion of infrastructure was highest in comparison to other categories of public spending over the same period. In addition, public spending on infrastructure was close to 10 per cent of GDP in 2007, which is considered as very high by international standards (World Bank, 2009). Generally, the growth performance of the economy and the emphasis of public spending on infrastructure discussed in Chapter 2 provide a rationale for an empirical assessment of the contribution of infrastructure to economic growth.

Chapter 3 examines whether infrastructure causes growth in Vietnamese provinces, after controlling for income level and a number of local attributes. Besides, this chapter also estimates the growth impact of cross-province infrastructure spillovers given that, as suggested by the literature, the spillover effect can arise from the connectivity characteristic of infrastructure, especially transport infrastructure. The panel dataset contains 61 provinces and covers the period 2000-2007. Empirical results obtained in this chapter are expected to contribute to understanding the importance of infrastructure improvement as an approach to economic development and to understanding how the provincial growth performance is influenced by the key aspects of economic reforms, particularly structural transformation and external trade, discussed in the previous chapter. In regard to methodology, growth regressions are estimated by the System-GMM which is aimed at controlling for the potential problems with endogeneity and serial correlation arising from the dynamic specification of the growth model. Furthermore, the System-GMM estimator can control for the endogeneity problem which may also arise from the inclusion of such an explanatory variable as the infrastructure variable. In addition, as Blundell and Bond (1999) and Bond *et al.* (2001) suggest, the System-GMM estimator is more efficient than the Difference-GMM estimator for a panel which is short in time dimension and includes persistent series.

The rest of Chapter 3 is organised as follows. Section 3.2 reviews the literature pertaining to the theoretical and empirical linkages between infrastructure and economic growth. Section 3.3 presents the model specification and estimation methods, describes variables added in the empirical models, and discusses issues related to the data used in the empirical analysis. Results are presented and discussed in Section 3.4. Finally, Section 3.5 summarises this chapter and draws policy implications.

3.2 LITERATURE REVIEW

Justification for public infrastructure investment comes from the theoretical belief of productive effects of infrastructure on economic development. Nevertheless, empirical evidence supporting this hypothesis is mixed. The first part of this section discusses the theoretical link between public capital and economic growth established in the endogenous growth theory, aimed at providing a theoretical justification for the role of the government, particularly public capital, in promoting economic growth. This is followed by a review of the empirical literature, focussing on empirical issues that could provide explanations for the conflicting results obtained in previous studies.

3.2.1 Public Capital and Endogenous Growth

Growth has been modelled following the two main approaches, namely, the neoclassical and endogenous growth modelling approaches. In essence, both neoclassical and endogenous growth models reach identical conclusions regarding the effect of public policy, in particular fiscal policy, on the level of output, whereas they are different in terms of their conclusions in regard to the growth effect of public capital. In the neoclassical growth models, fiscal policy can affect the level rather than the growth rate of output. More precisely, taxes and government expenditures that influence the investment incentive or the savings rate can affect the long-run growth only in the transition path to a steady state, whereas such exogenous factors as technological progress and population growth will cause the steady-state growth rate. The exogenous growth models are criticised due to their predication that, without the growth of exogenous factors, the economy will converge to the steady state with a zero growth rate. And, this zero growth rate can mainly be explained by the diminishing returns to capital. On the contrary, in the endogenous growth models, fiscal policy has a role to play in determining both the level of output and the steady-state growth rate. In particular, in these models, the positive constant growth rate of output is driven by government expenditures that raise the marginal product of capital to such an extent that diminishing marginal returns are permanently offset. The Barro (1990) model is considered as the first, and perhaps the simplest, example of endogenous growth models which encompass public capital into the production function (Irmen and Kuehnel, 2009). Later, extensions to the Barro (1990) model show that the growth effect of public capital is also dependent on the characteristics of public services, among other factors.

3.2.1.1 The Barro (1990) Model

To demonstrate how public services cause growth, Barro (1990) develops a model in which government expenditure is considered as an additional input to private production. In the Barro model, production exhibits constant returns to scale in private inputs and public services together but diminishing returns to scale in private inputs separately. As explained by Barro (1990), without a parallel expansion of public services, which are assumed to be a complementary input, production would involve decreasing returns to private inputs. The returns to scale are assumed to be constant. The production function for a representative household-producer is given as

$$y = \Phi(k, g) = k.\Phi\left(\frac{g}{k}\right) \tag{3.1}$$

where y is output per worker; Φ satisfies the conditions for positive and diminishing marginal products, so that $\Phi' > 0$ and $\Phi'' < 0$; k is private capital per worker; and g is public services derived by each household-producer. In addition, it is assumed that public services are not subject to congestion effects and are provided without user charges, meaning that public services are a pure public good.

Equation (3.1) can be written in the form of a Cobb-Douglas production function as

$$\frac{y}{k} = \Phi\left(\frac{g}{k}\right) = A.\left(\frac{g}{k}\right)^{\alpha} \tag{3.2}$$

where $0 < \alpha < 1$. And, the marginal product of private capital derived from equation (3.1) is

$$\frac{\partial y}{\partial k} = \Phi\left(\frac{g}{k}\right).\left(1 - \Phi'.\frac{g}{y}\right) = \Phi\left(\frac{g}{k}\right).(1 - \eta) \tag{3.3}$$

where, for a given value of k, η is the elasticity of y with respect to g and $0 < \eta < 1$. The marginal product, $\partial y/\partial k$, is calculated under the assumption that g is fixed and k varies in equation (3.1). In other words, changes in the quantity of capital k and output y do not lead to any changes in the amount of public services g. According to equation (3.3), an increase in g/y raises $\partial y/\partial k$, which finally raises growth rate, γ, specified in (3.6) and (3.7). Given that the term g/y also represents the size of government, Barro (1990) predicts that the growth-reducing effect dominates when the government is large, while the growth-enhancing effect dominates when the government size is small.

It is worth noting that Barro (1990) also emphasises two issues arising from the inclusion of public services as an additional input into the production function. The first issue is that the flow of public services does not necessarily correspond to government purchases. This happens when the government owns capital and, in the national accounts, the imputed rental income on public capital is omitted from the measure of current government purchases.

However, Barro (1990) suggests that we can assume that the government does no production and owns no capital. In this way, public services, which are purchased from private producers and then distributed to households, correspond to the input that is important for private production. Nevertheless, given that the production functions of the government and the private sector are the same, the results are supposed to be the same if the government does its own production using inputs purchased from the private sector, instead of purchasing only the final outputs produced by the private sector.

The second issue arises from the public-good characteristic of public services. If public services are non-rivalrous for the users, then the total amount of government purchases, rather than the amount in per-capita terms, would matter for each individual. As Barro (1990) notes, few public services have the non-rivalrous characteristic, and this is important for determining the desirable scale of public service provision. However, he also supposes that the private sector would not readily take the role of the government in providing the non-rivalrous services due to difficulties in relation to user charges. In particular, user charges would be undesirable either because of the non-rivalrous characteristic of the services or because of the low externalities associated with private production of the non-rivalrous services, and the provision of basic education is an example of this case. In this way, Barro (1990) implicitly provides recommendations regarding a productive role for the government to play for the development of the economy.

There should be the third important issue which arises from the assumption on whether public services are financed by a “lump-sum” tax or a “flat-rate” tax. Barro (1990) assumes that the government runs a balanced budget because it is unable to finance its deficits by issuing debt or run surpluses by a further accumulation of assets. Then, the budget constraint of the government is specified as

$$g = T = \tau y = \tau . k . \Phi\left(\frac{g}{k}\right) \tag{3.4}$$

where T is the total amount of taxes or government budget, and τ is tax rate.

In the presence of public services which are assumed to be financed by a "lump-sum" tax, the growth rate of consumption[13]

$$\gamma = \frac{\dot{c}}{c} = \frac{1}{\sigma} . (A - \rho) \tag{3.5}$$

is rewritten as

$$\gamma_L = \frac{\dot{c}}{c} = \frac{1}{\sigma} . \left[\Phi\left(\frac{g}{k}\right) . (1 - \eta) - \rho\right] \tag{3.6}$$

where c is per capita consumption; $\sigma > 0$ is the marginal utility elasticity; and $\rho > 0$ is the time preference, which is assumed to be constant.

However, if public services are financed by a "flat-rate" tax, equation (3.6) is modified as

$$\gamma = \frac{\dot{c}}{c} = \frac{1}{\sigma} . \left[(1 - \tau) . \Phi\left(\frac{g}{k}\right) . (1 - \eta) - \rho\right] \tag{3.7}$$

γ differs from γ_L by the presence of the term $1 - \tau$. If g and T are set to grow at the same rate as y, then g/k, η and, hence, γ will be constant. An increase in the tax rate, τ, reduces the

[13] which is obtained by substituting the marginal product of capital f' which is assumed to equal to the constant net marginal product of capital A, $f' = A$, into the consumption growth equation $\frac{\dot{c}}{c} = \frac{1}{\sigma} . (f' - \rho)$.

growth rate, γ; whereas, an increase in the ratio of public services to output, g/y, raises the growth rate, γ.

3.2.1.2 Extensions to the Barro (1990) Model

The majority of the extensions to the Barro (1990) model can broadly classified into two categories: those making changes to the characteristics of public services or government expenditures, and those making changes in the manner by which government expenditures affect the production of output (expenditures that encourage the additional accumulation of production inputs). The extensions belonging to the former category focuses on the degree to which public goods are subject to rivalry and excludability. For the latter, pubic capital is implicitly assumed to have pure public-good characteristic. Irmen and Kuehnel (2009) have provided a comprehensive survey of the extensions to the Barro (1990) model. Therefore, this section summarises the main extensions and then highlights the model recently developed by Hashimzade and Myles (2010) which shows that the production process also benefits from spatial infrastructure spillovers.

According to Irmen and Kuehnel (2009) the literature has identified two forms of congestion: (i) relative and (ii) absolute, or aggregate, congestion. Regarding the case of public services subject to relative congestion, the level of services distributed to a representative producer depends on its size relative to the aggregate of producers. In the case of absolute congestion, the level of public services distributed to the representative producer is decreasing in the aggregate usage. In their survey, Irmen and Kuehnel (2009) present three cases of public services subject to congestion: (i) public services subject to relative congestion but non-excludable; (ii) public services subject to relative congestion but excludable; and (iii) public services subject to absolute congestion, which are summarised below.

Firstly, the amount that the representative producer derives from the public services G subject to relative congestion is specified as $g = G\left(\frac{k}{K}\right)^{1-\varphi}$,[14] and the production function of the representative producer is written as,[15]

$$y = A\left(\frac{G}{K}\right)^{\alpha}\left(\frac{k}{K}\right)^{-\alpha\varphi} k \tag{3.8}$$

and the growth rate is given by[16]

$$\gamma = (1-\tau)(1-\alpha\varphi)(AN^{\alpha\varphi}\theta_G^{\alpha})^{\frac{1}{1-\alpha}} - \rho \tag{3.9}$$

According to equation (3.9) an increase in the congestion degree, that is, a decline in φ, has two effects on the growth rate, γ: (i) it augments the elasticity of output with respect to private capital, $1-\alpha\varphi$; and (ii) it weakens the scale effect through $N^{\varphi\alpha}$. Whether the former effect dominates the latter effect, or vice versa, depends on the number of household-producers and the congestion degree.

[14] where g is the amount derived by the representative producer from the public services G, k is the capital of the representative producer, G and K denote the aggregate public capital and the aggregate private capital, respectively, the ratio k/K measures the size of a representative producer relative to the economy, and the parameter $\varphi \in [0,1]$ measures the degree of relative congestion associated with the public services G. In this case of congestion, it is assumed that producers believe that an increase in k would raise their benefit from the publicly provided services and disregard the impact of their investment decision on $G$and K. If $\varphi = 0$, g increases if G grows faster than K. If $\varphi = 1$, then public services G are pure public goods. If $0 < \varphi < 1$, g increases if G and K grow at the same rate.

[15] where y is output per producer, and $A > 0$ is the constant total factor productivity.

[16] where γ is the growth rate, τ is the tax rate, $N > 1$ is the constant number of household-producers, $\theta_G \in (0,1)$ is the constant ratio of the aggregate public services G to the aggregate output Y, and ρ is the time preference.

Secondly, the amount that the representative producer derives from the public services subject to relative congestion but excludable, E, is specified as $e = E\left(\frac{k}{K}\right)^{1-\omega}$,[17] and the production function of the representative producer is modified as,

$$y = A\left(\frac{G}{K}\right)^{\alpha}\left(\frac{k}{K}\right)^{-\alpha\varphi}\left(\frac{E}{K}\right)^{\beta}\left(\frac{k}{K}\right)^{1-\omega}k \tag{3.10}$$

and the growth rate has the form[18]

$$\gamma = (1-\tau)(1-\alpha\varphi-\beta\omega)\left(AN^{\alpha\varphi+\beta\omega}\theta_G^{\alpha}\theta_E^{\beta}\right)^{\frac{1}{1-\alpha-\beta}} - \rho \tag{3.11}$$

As can be seen in equation (3.11), the government can influence the growth of the economy through the provision of two types of public services which have different public-good characteristics.

Thirdly, when public services are subject to absolute congestion, they are independent of the size of the representative producer. The amount that the representative producer derives from the public services subject to absolute congestion, P, is specified as $p = P(K)^{\vartheta-1}$. In this case of congestion, the representative producer's production ceases to exhibit constant returns to scale in both public capital and private capital and additional restrictive conditions are required for achieving the steady-state growth.

---

[17] where e is the amount derived by the representative producer from the public services E, and the parameter $\omega \in [0,1]$ measures the degree of relative congestion associated with the public services E. It is assumed that while the public services G must be financed by taxes, the public services E can be financed by user charges.

[18] where $\theta_E \in (0,1)$ is the constant ratio of the aggregate public services E to the aggregate output Y.

Barro (1990) assumes that the productive growth effects of different expenditure categories are identical, and therefore government expenditures can be aggregated into a single term.[19] In practice, the effect of transport infrastructure expenditure on the rate of growth is not likely to be identical to that of education expenditure or health expenditure. Moreover, the indirect effects of expenditures which encourage the additional accumulation of production factors seem to be ignored due to the assumption of identical effects. Examples of this type of expenditures can be R&D, education or health expenditures. Devarajan *et al.* (1996) find evidence against the assumption on the homogeneous growth effects of different categories of government expenditures. Theoretically, they add two categories of government expenditures into the same production function which is specified as[20]

$$y = f(k, g_1, g_2) = \left(\alpha k^{-\xi} + \beta g_1^{-\xi} + \mu g_2^{-\xi}\right)^{-1/\xi}$$
$$\alpha > 0, \beta \geq 0, \mu \geq 0, \alpha + \beta + \mu = 1, \xi \geq -1 \tag{3.12}$$

It is assumed that the government runs a balanced budget and the two categories of government expenditures are financed a "flat-rate" income tax, $\tau y = g_1 + g_2$.[21] This implies that the levels of these government expenditures are determined by default, while the share of each expenditure category is given by $g_1 = \psi\tau y$ and $g_2 = (1-\psi)\tau y$. Provided that τ and ψ are taken as given, the representative agent chooses consumption, c, and capital, k, to maximise his welfare subject to his budget constraint. Then, the equation for the steady-state growth rate of consumption has the form[22]

[19] Barro (1990) also examines the case of government consumption which affects the household utility, but does have any effect on the production.

[20] where y and k denote output and capital of the representative agent, respectively; and g_1 and g_2 are the two categories of public services.

[21] In addition, τ and g/y are assumed to be constant.

[22] where γ is the growth rate, c is the consumption of the representative agent, τ is the tax rate, and ρ is the time preference.

$$\gamma = \frac{\dot{c}}{c} = \frac{1}{\sigma}.\left\{\alpha(1-\tau)\left\{\alpha\tau^{\xi}/\left[\tau^{\xi} - \beta\psi^{-\xi} - \mu(1-\psi)^{-\xi}\right]\right\}^{-(1+\xi)/\xi} - \rho\right\} \tag{3.13}$$

The component g_1 can be defined as productive if

$$\frac{\partial\gamma}{\partial\psi} = \frac{\alpha(1-\tau)(1+\xi)\left(\alpha\tau^{\xi}\right)^{-(1+\xi)/\xi}\left[\beta\psi^{-(1+\xi)} - \mu(1-\psi)^{-(1+\xi)}\right]}{\sigma[\tau^{\xi} - \beta\psi^{-\xi} - \mu(1-\psi)^{-\xi}]^{-1/\xi}} > 0 \tag{3.14}$$

However, Devarajan *et al.* (1996) also show that whether a change in the expenditure composition could raise the growth rate depends not only on the productivity of the two categories but also on their initial shares in total government expenditures. Specifically, a shift towards an increase in the share of the productive expenditure g_1, whose initial share ψ is "too high", might not raise the growth rate γ.

- In the Barro (1990) model, public capital is assumed to be a flow variable and, therefore, the economy is in a position of steady-state growth in which all quantities, namely, consumption per capita, private capital per capita, public capital per capita, and national product per capita, grow at the same rate. Futagami *et al.* (1993) propose an endogenous growth model in which public capital and private capital are assumed to be stock variables. Consequently, there are transitional dynamics in that model. Their main conclusions can be are summarised as follows.
- A unique steady-growth equilibrium is achieved when $\frac{[(1-\tau)(\sigma+\bar{\eta}-1)]\bar{\Phi}}{\sigma} + \frac{\rho}{\sigma}$ is positive;[23]
- There is a unique stable path converging to the steady-growth equilibrium when there exists the unique steady-growth equilibrium;

23 where τ denotes a "flat-rate" income tax, ρ denotes the time preference, σ denotes the inverse of the intertemporal substitution elasticity, η denotes the output elasticity of public capital, $\bar{\eta} \equiv \eta(\frac{\tau}{1-\tau})$, and $\bar{\Phi} \equiv \Phi(\frac{\tau}{1-\tau})$.

- If the output elasticity of public capital, η, is constant, the steady-growth equilibrium reaches its maximum growth rate when the income tax rate equals the output elasticity of public capital, $\tau = \eta$;
- If η is constant and given that $\eta \leq \tau$, the dynamics of the economy are given by $\frac{\partial y(0,\tau)}{\partial \tau} \geq 0$ when $\sigma + \eta \leq 1$ and by $\frac{\partial y(0,\tau)}{\partial \tau} \leq 0$ when $\sigma + \eta \geq 1$;
- If η is constant and given that $\sigma = 1$, the optimal income tax rate is lower than the rate that maximises the growth rate of the economy. This conclusion has important implications for the infrastructure policy aimed at promoting growth since public infrastructure investment is assumed to be financed by taxation.

The final extension to the Barro (1990) model worth being discussed here is the model developed by Hashimzade and Myles (2011) in which there are two countries, namely, the "home" country and the "foreign" country. It is assumed that, in each country the government runs a balanced budget, and government expenditure is financed by a tax levied on private capital. Furthermore, public services are assumed to be pure public goods. The aggregate production function of the "home" country has the following Cobb-Douglas form

$$Y = AK^{\alpha}(G^{1-\nu}J^{\nu})^{1-\alpha} \tag{3.15}$$

where Y is output; A is a positive constant; K is private capital; G is public infrastructure investment in the "home" country; and J is global infrastructure. J equals to $G + \bar{G}$, which comprises the public investment in infrastructure in the "home" country, G, and the public investment in infrastructure in the "foreign" country, $\bar{G}$. The spatial spillover effect of infrastructure stems from the inclusion of the term $\bar{G}$.

The growth rate in the Barro-typed one-country model, which has the form

$$\frac{\dot{c}}{c} = \gamma = \beta\left[\alpha A\left(\frac{G}{K}\right)^{1-\alpha} + 1 - \delta - \tau\right] \tag{3.16}$$

is rewritten, in the context of the two-country model, as

$$\frac{\dot{c}}{c} = 1 + \gamma = \beta\left[\alpha A\left(\frac{G}{K}\right)^{1-\alpha} + (1 + \frac{\bar{G}}{G})^{\nu(1-\alpha)} + 1 - \delta - \tau\right] \tag{3.17}$$

As can be seen in equation (3.17), an increase in public infrastructure in the "home" country and an increase in infrastructure in the "foreign" country raises the growth rate in the "home" country. Given that the production function of the "foreign" country is specified in the same way to that of the "home" country and that the parameters are assumed to be the same for both countries, an increase in infrastructure in one country promotes the growth in all countries.

In the presence of a positive growth impact of spatial infrastructure spillovers, Hashimzade and Myles (2010) also show that policy coordination between the "home" and the "foreign" countries could make public infrastructure provision more efficient. As they argue, since infrastructure is assumed to be financed by taxes, the growth effect of infrastructure and the cross-country spillovers of infrastructure raise the importance of both the tax policy of each country and the taxation effects distributed across countries. Also, they suggest a role for a supra-national organisation to play in coordinating the policy decisions of individual countries for an increased welfare. It is worth noting that this model is developed in the context of the European Union. Then, to some extent, the model can serve as a theoretical foundation for an empirical analysis on the cross-region, or the cross-province, spillover effect of infrastructure at the sub-national level.

3.2.2 Empirical Literature on the Relationship between Infrastructure and Economic Growth

In the empirical literature, there exists an on-going debate regarding whether infrastructure causes economic growth. Straub (2008) and, more recently, Pereira and Andraz (2010) provide comprehensive surveys of empirical evidence in this line of research. This section begins with a summary of several figures generalised from previous surveys which could be a general illustration of the current state of the empirical literature. This is followed by a discussion on several econometric issues associated with the estimation of the growth impact of infrastructure, drawbacks of monetary and physical infrastructure measures, geographical and sectoral dimension of growth analysis, and spillover effects of transport infrastructure. To some extent, the empirical issues discussed in this section could provide explanations for the conflicting results obtained in previous studies.

3.2.2.1 Some of Previous Surveys of Empirical Evidence

Straub (2008) surveys the empirical literature, basing on 140 empirical specifications of 64 papers published during the period 1989-2007. The level and the growth rate of output, and productivity, have been the most frequently used dependent variables. In regard to the geographical dimension of empirical analysis, most specifications investigate the impact of infrastructure at the national level. Specifically, the proportions of specifications using aggregate data, sub-national data, sectoral data, and micro data are about 58 per cent, 25 per cent, 6 per cent, and 11 per cent, respectively. Concerning infrastructure proxies, between 1989 and 1999, the proportion of papers using monetary indicators is 72 per cent against 28 per cent using physical indicators. The reversed figures can be seen over the period 2000-

2007: 24 per cent using monetary indicators but 76 per cent using physical indicators. With respect to the sample, in the period 1989-1999, only 29 per cent of specifications used developing country data; while the figures for developed country samples and mixed developed/developing country samples are 51 per cent and 20 per cent, respectively. For the period 2000-2007, the figures turn out to be 47 per cent, 18 per cent and 35 per cent, respectively. Regarding the methodology, macro-econometric techniques play a dominant role in this field of research, whereas micro-level specifications, using either firm-level or household data, account for only about 18 per cent of the sample. Among the papers adopting macro-econometric techniques, specifications following the production function approach represent 67 per cent of the sample. This is followed by cross-country regressions with 21 per cent, cost function estimations with 9 per cent and growth accounting techniques with 3 per cent. In relation to empirical findings, about 63 per cent of the specifications obtain significantly positive results regarding the impacts of infrastructure on development outcomes, about 31 per cent show statistically insignificant results, and only about 6 per cent find a negative and statistically significant relationship.

A recent survey of Pereira and Andraz (2010) is based on 155 papers between 1978 and 2010. In regard to the sample of countries, only 2 per cent of papers use developing country data and 6 per cent use mixed developed/developing samples. Concerning the level of data used in the empirical estimations, the proportions of studies using aggregate level data, regional level data, and sectoral level data, are 49 per cent, 36 per cent, and 15 per cent, respectively. In terms of estimation approaches, papers following the production function approach comprise 55 per cent of the survey's sample, followed in frequency by those following VAR approach (28 per cent) and profit or cost function approach (17 per cent). Generally, the production function approach has been the most popular approach to growth analysis, and empirical evidence on the growth impact of infrastructure in developing countries remains limited.[24]

Disaggregated figures presented in Straub (2008) appear to indicate that the distribution of empirical results vary according to different categories of samples, dependent variables, infrastructure proxies, and estimation techniques. Firstly, the figure of positive results is higher for the group of specifications using developing country data, compared to the groups using developed-country samples and mixed samples. Findings obtained by studies using mixed sample are often inconclusive, and this can be attributed to the fact that there exist unobservable heterogeneous characteristics among cross-sectional units. Secondly, the figure of positive results is higher at a more disaggregated level of the analyses. This, according to Straub (2008), can be seen by comparing between country-level studies and firm-level studies. Thirdly, only about 54 per cent of specifications using either output or output growth as the dependent variable and about 67 per cent of specifications for productivity produce positive results. Fourthly, specifications using physical proxies for infrastructure are much more successful in obtaining positive results in relation to the impact of infrastructure on development outcomes, in comparison to those using monetary infrastructure measures. Finally, the distribution of results across specifications following the production function approach is slightly similar to the overall distribution.

3.2.2.2 Econometric Issues

Aschauer (1989) opens the debate on this line of research. His empirical analysis is based on the production function approach to estimate the impact of infrastructure on economic growth in the U.S., and finds an output elasticity of 0.39 for total public capital and an output

[24] These figures are calculated using the list of surveyed studies presented in Pereira and Andraz (2010).

elasticity of 0.24 for core infrastructure. Importantly, those elasticises are larger than the elasticity estimated for private capital. Subsequent studies confirm the results obtained by Aschauer (1989) at the U.S. national level (e.g., Munnell, 1990; Nadiri and Mamuneas, 1994). Nevertheless, some studies find smaller output elasticities of public capital, and even statistically insignificant coefficient on the public capital variable, at the U.S. state level (e.g., Hulten and Schwab, 1991; García-Milà and McGuire, 1992; Morrison and Schwartz, 1996). On econometric grounds, the large output elasticities of public capital reported in the early studies were criticised to be the results of estimation biases arising from the uncontrolled problems associated with stochastic trends, the uncontrolled endogeneity of public capital, and the misspecification due to omitted variables.[25] That econometric problems are not equally solved in all studies provides one explanation for the fact that the empirical literature contains mixed results on the contribution of infrastructure to economic growth.

Reverse causality seems to be the major problem facing studies based on the production function approach. Much of empirical works hypothesise that economic growth is a function of infrastructure. However, the relationship between infrastructure and economic growth can be bi-directional. Specifically, while infrastructure can be a driver of economic growth, economic growth can also determine the demand for and hence the supply of infrastructure. Estimation biases may be caused by the failure to account for the potential endogeneity of infrastructure which could be done by using econometric techniques that can disentangle the causality between the infrastructure variable and the growth variable, such as the Instrumental Variables (IV) and the Generalised Method of Moments (GMM). Particularly in the growth analysis with the presence of the lagged dependent variable, the GMM estimator appears to be the most viable option.

[25] The critics to the empirical results obtained by Aschauer (1989), among others, are discussed in Gramlich (1994) and Pereira and Andraz (2010) for example.

A meaningful analysis of the relationship between infrastructure and economic growth requires that the potential growth impacts of other factors should also be estimated so as to control for estimation biases due to omitted variables. In fact, almost all studies add other variables, beyond infrastructure, to control specific characteristics of the countries or regions observed. However, the list of those variables is endless, and the set of explanatory variables varies significantly across studies. According to Sala-i-Martin (1997) the variables potentially explaining growth could be traditional economic variables, such as labour and physical capital, and a broader range of economic variables, such as human capital, R&D investment and public capital, and non-economic variables, such as social capital and institutions. As Levine and Renelt (1992, p.943) argue, "... almost all identified relationships are very sensitive to slight alternations in the conditioning set of variables." In this sense, the lack of a uniform set of explanatory variables implies that different specifications control for different things and therefore may yield very different results. This provides another explanation for the mixed results in regard to the growth impact of infrastructure.

3.2.2.3 Measurement of Infrastructure

Infrastructure can be measured by either monetary or physical approaches. Both approaches have their own drawbacks. What is important is that the mixed results on the relationship between infrastructure and economic growth could also be due to differences between empirical studies with regards to proxies employed for infrastructure.

As mentioned above, the empirical literature has gradually shifted from using monetary indicators to using physical indicators of infrastructure over the past decades (Straub, 2008). However, one drawback associated with the majority of physical indicators arises from the fact that they are unable to take into account of the quality of infrastructure. The quality of roads, for example, can vary across countries, and especially across sub-national regions in developing and least developed countries. In practice, the availability of data on infrastructure quality is even more limited than that of quantity data (Straub, 2008).

Beyond physical indicators of infrastructure, there are several attempts at constructing a single index capturing the characteristics of a variety of physical infrastructure categories. For instance, Calderon and Serven (2004) apply principal component analysis to disaggregate infrastructure indicators so as to construct an infrastructure stock index, which is based on such quantity indicators as main telephone lines, electricity generating capacity, and road length, and an infrastructure quality index, which is based on such quality indicators as waiting time for telephone main lines, percentage of transmission and distribution losses in electricity production, and ratio of paved roads to total roads. However, the construction of these indexes requires the availability of data for a number of infrastructure sectors.

For the monetary approach, one proxy for infrastructure can be the monetary value obtained by the perpetual inventory method of measuring the value of capital stock. Public infrastructure spending is another monetary proxy for infrastructure. However, there are several reasons why monetary figures cannot be good proxies for infrastructure.

According to Biehl (1991), since construction costs might vary across areas according to their specific landscapes, the monetary figures could not reflect the physical capacity of infrastructure. Furthermore, Romp and de Haan (2005) note that the use of monetary figures is likely to be inappropriate in the case where infrastructure is built in networks. Another drawback of the monetary approach is, as pointed out by Pritchett (1996), that the final costs and the effective value of publicly provided infrastructure are often disconnected from each other mostly due to inefficiencies or weaknesses associated with governance and institutions.

Tanzi and Davoodi (1997) argue that corruption could intensify the limitation of using public spending as a measure of infrastructure because it raises the costs of investment while decreasing quality and productivity of public services. It is worth noting that, as claimed by the World Bank (2006, p.82), "opportunities for corruption could arise at most stages of the infrastructure project cycle".

Some other reasons for the fact that the monetary figures do not reflect the effective infrastructure stocks and the level of infrastructure services arise from the perspective of fiscal policy making. On the financing side, in the case that taxation is a primary financing source of public infrastructure investment, the estimation of the effect of infrastructure expenditure could not be meaningful without a control of the countervailing effect of taxes on the economy. On the spending side, allocating public funds to infrastructure investment may reduce funds available for provision of other public services, such as health and education, and other growth-promoting programmes. Therefore, the interpretation of the effect of public infrastructure investment, which was not disentangled from the confounding effects of taxes and other categories of government expenditures, could be misleading.

3.2.2.4 Geographical and Sectoral Dimension of Growth Analysis

The empirical literature has been dominated by developed country studies. The question is whether policy implications are similar for countries at different stages of development. What is important is that countries at different stages of development face with different developmental issues and they are also at different levels of infrastructure quality. As noted by Straub (2008), the fact that findings obtained by studies using mixed samples of developed and developing countries are more often inconclusive can be in part explained by the unobservable specific characteristics of these two broad groups of countries.

Particularly at the sub-national level, regional differences have an important role to play in regard to their influence on the relations between public infrastructure and economic growth. Concerning the regional development policy with infrastructure investment, Hansen (1965) disaggregates public capital into two groups (namely, economic overhead capital - EOC - which facilitates the productive activities and the movement of economic goods and social overhead capital - SOC - which enhances human capital and provides social services) and regions into three groups (namely, congested, intermediate, and lagging regions), and hypothesises that the development outcomes of infrastructure provision are largely dependent on the specific conditions of a region. The central idea of the Hansen hypothesis is that different types of infrastructure have different development outcomes, and the development outcomes of infrastructure provision are largely dependent on specific conditions of a region. In particular, a greater growth impact of EOC could be seen in intermediate regions rather than in either congested or lagging regions whereas lagging regions would benefit from SOC. In addition, the regional development literature considers regional infrastructure policy as a solution to the development gap between regions, and emphasises that the impact of infrastructure depends on the presence of other favourable conditions in the recipient region, such as the availability of skilled workers, the degree of agglomeration, and the structure of the economy (e.g., Hansen, 1965; Nijkamp, 1986; Vickerman, 1991). In this sense, public provision of infrastructure appears to be necessary but not sufficient to generate any desired development outcomes, and factors that characterise the differences between countries or sub-national regions in various terms can influence how public infrastructure affects economic growth.

A recent country-level example is the study of Straub and Terada-Hagiwara (2010) on the growth impact of infrastructure in developing Asian countries. The authors find a positive impact of infrastructure endowment on GDP growth. When the interaction terms between income levels, proxied by income-based dummy variables for specific groups of developing countries, and their infrastructure indicators are added into regressions, the estimation shows that the net effect of infrastructure is lower in the low- and middle-income groups than in the high-income group. To explain this finding, the authors emphasise the importance of regional integration, which has been established among the high-income countries, to enhance the beneficial effect of infrastructure investment in those countries. Another reason is that the high-income group also has a more favourable institutional environment which in turn enhances the positive growth impact of infrastructure.

At the sub-national level, Majumder (2004) test the validity of the Hansen hypothesis in the context of India. In particular, the author classifies Indian districts into three groups according to their development levels, and examines whether the impacts of different categories of infrastructure vary with different groups of districts. The results show that social infrastructure is more important for growth in lagging districts in comparison to other categorises of infrastructure, whereas physical and financial infrastructure are more important for the intermediate and advanced groups. As the author argues, these results support the Hansen hypothesis, hence suggesting that infrastructure investment should be designed differently for different development-level based groups of districts to maximise the beneficial effects.[26]

From the sectoral dimension, there are studies showing that public capital affects industries differently and industries react differently to different components of infrastructure. This implies that the results found at the aggregate level could not show industry-level effects of

---

[26] The Hansen's hypothesis has been empirically tested in several contexts such as output, productivity and employment. The most frequently used approach involves classifying regions into several broad categories (for example, congested regions, intermediate regions, and lagging regions) and comparing results from these separate regressions. However, whether the Hansen hypothesis holds in the empirical context has been inconclusive.

infrastructure. According to the survey of Pereira and Andraz (2010), the majority of sectoral growth analyses find that the manufacturing sector is likely to benefit from infrastructure investment, while agriculture, which is traditionally considered as a declining sector, does not seem to benefit much.

3.2.2.5 Spatial Spillovers from Infrastructure

As argued by Mikelbank and Jackson (2000), the variations of empirical findings from no role to a substantial role for infrastructure could be in part due to the fact that some studies explicitly controlled for spatial externalities of infrastructure while others did not. In essence, the spatial spillover effect is explained by network characteristics of physical infrastructure, especially transport infrastructure. Some studies attribute their findings on a lower productive effect of infrastructure at the geographically disaggregated level to the existence of spatial spillovers from infrastructure (e.g., Hulten and Schwab, 1991; García-Milà and McGuire, 1992; Boarnet, 1998; Perara and Roca-Sagalés, 2003) and therefore argue that the spatial externalities of infrastructure could be more pronounced at the sub-national level.

Regarding the nature of spatial spillovers from infrastructure, there are two conflicting hypotheses. On the one hand, infrastructure may have a positive effect not only in a particular area where it is installed but also in contiguous areas or even other distant areas. In particular, positive spillovers can be caused by productivity leakages because of the connectivity characteristics of transport facilities (Munnell, 1992). On the other hand, there exist potentially negative spillovers from infrastructure. This is because infrastructure improved in a particular area enhances the comparative advantage of that area over the others and therefore could attract production inputs from other areas where might face with a decrease in output or productivity. More precisely, spillovers of transport infrastructure may arise from the migration of production factors, such as labour and physical capital, from areas with relatively low levels of infrastructure endowment to those with well-developed infrastructure, and such the spillovers are positive for regions of destination, but negative for regions of origin. As verified by Boarnet (1998), among others, such the migration of production inputs results in output gains in areas with well-developed transport capital stocks and output losses elsewhere. However, that a particular area is affected more or less by infrastructure installed in other areas also depends on several specific characteristics of that area such as economic structure and/or outward connections (Pereira and Andraz, 2004).

The production function with the infrastructure spillover variable has the Cobb-Douglas form, as specified in Boarnet (1998),

$$\ln Y_i = \alpha_0 + \alpha \ln K_i + \beta \ln L_i + \gamma \ln G_i + \delta \sum_{j=1}^{N} w_{i,j} \ln G_j \tag{3.18}$$

where Y is output, K is private capital, L is labour, G is the stock of public capital, i denotes a particular area, j denotes neighbouring areas, $w_{i,j}$ is the spatial weights matrix, N is the number of neighbouring areas and δ is the parameter measuring the effect of public capital stock of neighbouring areas on the output of area i. The statistically significant and positive (or, negative) estimate of the parameter δ indicates a positive (or, negative) spatial spillover effect.

In the empirical implementation, the most frequently-used method of constructing the matrix $w_{i,j}$ is based on a simple contiguity definition where $w_{i,j} = 1$ if i and j share a common boundary, and zero otherwise. The group of j can be modified to include further other nearest neighbours provided that the effect of infrastructure in one area could extend far beyond its bordering areas. However, such the simple spatial weights matrix ignores the nature and degree of the spatial interactions of economic activities that actually happen. An alternative is the matrix $w_{i,j}$ in which the elements are set equal to the inverse of the distance between areas, thereby emphasising a distance decay effect. Another alternative is the matrix $w_{i,j}$ which is constructed by using such elements as population density and income level so as to capture the possibility of negative spillovers from infrastructure due to the competition for mobile production factors between areas having similarity or dissimilarity in regards to their socioeconomic characteristics.

Of major importance is the fact that the spatial weights matrix defines how the neighbour relations have to be controlled for. Specifically, the matrix is used to evaluate the covariance of characteristics across areas and it contains information about the relative dependence between areas in the sample. In relation to the construction of the matrix, it is suggested that the matrix must not contain any of the exogenous or endogenous variables used in the regression to ensure that the empirical model not become highly non-linear, and this could be the reason why the most widely used weighting methods rely on contiguity and distance between areas (e.g., Anselin and Bera, 1998; Abreu *et al.*, 2005).

The existence of spatial infrastructure spillovers has been investigated empirically at a variety of geographical level. Findings have so far been mixed, however. Studies using the United States' data find either positive effect (e.g., Dalenberg *et al.*, 1998), or negative effect (e.g., Haughwout, 1999; Chandra and Thompson, 2000; Chalermpong, 2004; Cohen and Paul, 2004) or even no evidence (e.g., Holtz-Eakin and Schwartz, 1995; Kelejian and Robinson, 1997) of transport infrastructure spillovers across spatial units (i.e., states, counties, etc.). In the context of Spain, some studies do not find spatial spillovers of highways across Spanish provinces (e.g., Alvarez *et al.*, 2006) while others show evidence of either positive (e.g., Mas *et al.*, 1996; Gil et al., 1997; Avilés *et al.*, 2003; Pereira and Roca-Sagalés, 2003; Lanzas and Martínez, 2003; Ezcurra *et al.*, 2005; Cantos *et al.*, 2005) or negative (e.g., Moreno and López-Bazo, 2005; Delgado and Álvarez, 2007) spatial spillovers. There are a small number of studies using data of other countries, such as Liu *et al.* (2007) for Chinese cities, which finds a positive spillover effect, and Bronzini and Piselli (2008) for Italian regions, which also shows positive evidence. In regard to the possibility of cross-country spillovers from public capital, Owyong and Thangavelu (2001) specify a growth model for Canada in which the U.S. public capital was introduced as an exogenous variable. They find positive spillovers from the U.S. public capital to the productivity of the Canadian economy, hence concluding that the cross-country spillover effect could exist in the case of countries that are integrated and allow a free movement of goods as well as production factors.

In addition to the discussed-above issues associated with econometric techniques, infrastructure measures and the set of explanatory variables controlling specific characteristics, one explanation for the mixed empirical findings on the existence of spatial infrastructure spillover effect can be derived from differences in geographical scope of the spatial units observed. For instance, while Moreno and López-Bazo (2007) find negative spillovers from transport infrastructure across Spanish provinces, they compare their findings with those findings on positive spillovers for larger regions of Spain in other studies, such as Mas *et al.* (1996), Pereira and Roca-Sagalés (2003) and Avilés *et al.* (2003), and suggest that the net positive spillover effect might appear at a broader geographical level. For the case of the United States, Pereira and Andraz (2004) believe that the larger size of the regions in their study (i.e. states) masks the negative effects found in other studies at a smaller regional scale such as the study of Boarnet (1998) which analyses spillovers at the county level. Pereira and Andraz (2010) also provide a survey of empirical findings on transport infrastructure spillovers and conclude that the net positive spillovers might be identified at a broader geographical level while, across a relatively short distance, the negative spillovers might overwhelm the positive ones.

3.3 EMPIRICAL MODELS, VARIABLES AND DATA

3.3.1 Model Specification and Estimation Method

There is a substantial literature on the "Barro-regressions" after an influential early study by Barro (1990). The Barro-type growth regression has the form

$$g_{i,t} = \alpha + \beta \ln y_{i,t-1} + \gamma \mathrm{X}_{i,t} + \varepsilon_{i,t} \tag{3.19}$$

where g denotes the growth rate of income per capita, y denotes the level of income per capita, and X is a vector of variables characterising each cross-section i which are assumed to be of influence for the steady state level.

In the present study, the growth model is specified as

$$\begin{aligned} \ln GDPPC_{i,t} - \ln GDPPC_{i,t-1} \\ = \alpha + \beta \ln GDPPC_{i,t-1} + \gamma \ln ROAD_{i,t} + \delta_X X_{i,t} + \mu_i + \varphi_t + \varepsilon_{i,t} \end{aligned} \tag{3.20}$$

where $\ln GDPPC_{i,t} - \ln GDPPC_{i,t-1}$ is the measure of the growth of GDP per capita in province i in year t; $\ln ROAD_{i,t}$ denotes the density of the road network, measured by total road length per 1,000 km^2, in province i in year t; $X_{i,t}$ is a vector of other explanatory variables measured for province i in year t; μ_i is the unobserved province-specific factors; φ_t is the unobserved time-specific factors; and $\varepsilon_{i,t}$ is the error terms.

Based on the proposition of the potential existence of cross-region infrastructure spillovers, as reviewed in the previous section, the present study also examines the growth impact of transport infrastructure by estimating the following model which incorporates both own-province and neighbouring provinces' transport infrastructure

$$\ln GDPPC_{i,t} - \ln GDPPC_{i,t-1} = \alpha + \beta \ln GDPPC_{i,t-1} + \gamma \ln ROAD_{i,t} + \delta_X X_{i,t} + \Omega \ln w_{i,j} ROAD_{i,t} + \mu_i + \varphi_t + \varepsilon_{i,t} \quad (3.21)$$

As it is widely acknowledged, the traditional cross-sectional growth regressions can carry such econometric problems as measurement errors, endogenous variables, and omitted variables. The panel data approach can be applied for growth regressions to alleviate the problem of omitted variables by taking into account of cross-section-specific and time-specific effects (Islam, 1995). Furthermore, the instrument-variable estimation method of Two-Stage Least Squares can be applied to tackle with the problem of endogeneity. However, Nickell (1981) points out the "dynamic panel bias" which arises from the correlation between the lagged dependent variable and the error terms in the dynamic growth model. Then, Arellano and Bond (1991) propose the application of the Difference Generalised Method of Moments (Difference-GMM) estimator for the dynamic panel data model. Later, Arellano and Bover (1995) and Blundell and Bond (1999) develop the System Generalised Method of Moments (System-GMM) estimator by introducing an additional stationarity restriction to the Difference-GMM estimator.

Blundell and Bond (1999) and Bond *et al.* (2001) suggest that the System-GMM estimator is more efficient than the Difference-GMM estimator in the case that the panel is short in time and contains persistent time series. Accordingly, if the number of periods is small, the Difference-GMM estimates are biased since they tend to fall below the corresponding within-group estimates in a fixed effect model. Furthermore, if the series are persistent over time, the Difference-GMM estimator is inefficient because the lagged levels of the series provide weak instruments. In the present study, given the nature of the panel dataset, the System-GMM estimator is employed for all growth regressions.

The consistency of the GMM estimator depends on the validity of the instruments which is examined by two specification tests proposed by Arellano and Bond (1991), Arellano and Bover (1995) and Roodman (2006). The first is the Hansen test of overidentifying restrictions which considers the overall validity of the instruments. Accordingly, if the null hypothesis that the instruments are not correlated with the residuals cannot be rejected, the instruments are valid. With respect to the second test of no serial correlation in the residuals, the validity of the instruments requires that the null hypothesis of no second-order serial correlation cannot be rejected while the null hypothesis of no first-order serial correlation is rejected. After all, the specification of the model is supported as long as the validity of the instruments is confirmed.

3.3.2 Variables

Road density, i.e. road length per 1,000 km^2 of the provincial area, is used as a proxy for the availability of transport infrastructure. The present study assumes that an increased road density is an indicator of the road network improvement. Another assumption is that the quality of roads is homogenous at the province level. A positive sign is expected for the coefficient on the variable $ROAD$.

As suggested in the growth literature, macroeconomic instability and, in particular, a high inflation rate tends to affect the growth performance negatively. Easterly (2003) analyses on a set of six variables capturing several dimensions of public policies: inflation, budget balance, real overvaluation, black market premium, financial depth, and trade openness. It is worth noting that the study of Easterly (2003) is carried out at the country level. Some of those policy variables might not have a significant role to play in explaining growth performance at the sub-national level, at least in the context of Vietnam which is administratively divided into a proportionate number of small-sized units and where the economic centres, and hence economic activities, have been concentrated only in Hanoi and Ho Chi Minh City. On the other side, there are no province-level data for those variables. Therefore, the present study follows Easterly (2003) to examine the growth impact of macroeconomic policy, but focussing on the impact of inflation. Indeed, the inclusion of the inflation variable, $INFLATION$, measured by a change in the logarithm of the consumer price index, may suffer from the criticism that inflation may not be helpful to explain the provincial growth performance. However, the present study hypothesises that an increase in the price index generates adverse effects not only on macroeconomic stability but on the welfare of people residing in a particular province and therefore expects a negative sign for the coefficient on this variable.

Structural transformation has been one of the most striking aspects of the transition economy of Vietnam. In particular, a transition of economic activity from agriculture to non-agricultural sectors is expected to lead to faster growth, since it implies a structural shift from lower- to higher-productivity sectors. The present study uses the change of the proportion of agriculture in GDP, $\Delta AGRGDP$, to capture the transformation in the sectoral composition of GDP. The agriculture sector, in a broad sense, includes the agriculture, forestry and fishery industries. A negative sign is expected for the coefficient on this variable.

Another important aspect of economic reforms in Vietnam has been the process of structural transformation in terms of ownership. As discussed in Chapter 2, while the efficiency and profitability of SOEs generally remained limited, the private sector grew rapidly in response to market opportunities over the period 2000-2007 in which the economy also achieved a high growth rate of GDP. The present study uses the proportion of physical capital stock owned by SOEs in total physical capital stock, $SOECAP$, as a proxy for the ownership structure. Alternatively, the ownership structure can be proxied by the proportion of SOE employment in total employment, $SOEEMP$, or the proportion of SOEs in total number of enterprises, $SOEENT$. The sign of the coefficient on each of these ownership structure variables is expected to be negative.

With respect to the growth impact of external trade, the present study examines the performance of the following variables in the growth regressions: the ratio of exports to GDP, $EXPGDP$, and the ratio of imports to GDP, $IMPGDP$. These variables can also be considered as proxies for the progress in implementation of the opening-up policies which have been one of the core aspects of economic reforms in Vietnam. However, the growth impact of exports and imports seem to be ambiguous as long as the empirical literature on the relationship between external trade and economic growth remains controversial.

The change of in the proportion of agriculture in GDP captures a transformation in the sectoral structure of the economy. Urbanisation presents a factor that is fundamental to "a multidimensional structural transformation" (Annez and Buckley, 2008, p.1) that is necessary to sustain growth in developing countries. In the present study, the degree of urbanisation, $URBAN$, is measured by the proportion of urban population to total population. One rationale for the inclusion of the urbanisation variable is to examine the externalities of the concentration of people in urban centres. However, as presented in Chapter 2, the phenomenon of over-urbanisation has also been identified in the most developed cities of Vietnam. Therefore, the net effect of urbanisation on growth might be ambiguous.[27]

Given that a greater endowment of infrastructure contributes to a particular province's economic growth, a related policy issue is whether the beneficial impact of infrastructure could also spill over into its neighbouring provinces. In particular, road infrastructure is essentially a network and can thus exert an effect beyond where it resides in space. The present study includes the average road density in the neighbouring provinces as an additional explanatory variable to estimate the potential existence of cross-province spillover effect of transport infrastructure. The literature reviewed above suggests that growth performance of a particular area can also be influenced either positively or negatively by infrastructure installed in its neighbouring areas. Then, the sign of the coefficient on the road spillover variable, $wROAD$, is left to be empirically determined.

The natural criterion for considering cross-province infrastructure spillovers in this context is adjacency, since provinces that share borders are assumed to be most likely to experience spillovers from each other's infrastructure, than provinces not linked, or those at one remove from the immediate vicinity. Two different spatial weights matrices are used alternatively to calculate the variable $wROAD$. The first spatial weights matrix is a binary contiguity-based measure, w_{cont}, where a weight of 1 is assigned if the two provinces share a common boundary, and zero otherwise. The second spatial weights matrix is based on the inverse of geographical distance between the centres of the two provinces sharing a border. While the contiguity- and distance-based spatial weights matrix, w_{dist}, captures the spatial spillover

[27] According to GSO (2011b), urban people have better housing conditions, greater access to basic amenities, and more chances to attain higher education and more chances to obtain professional jobs. On the other side, excessive concentration of population in a few cities combined with the current inadequate endowment of urban infrastructure lead to the phenomenon of over-urbanisation. This in turn implies congestion in using basic amenities as well as other public services and hence the beneficial effect of urbanisation could be hindered.

effect which is assumed to diminish with distance, it could also account for the role of geographical proximity in the competition for production inputs that could be mobile across provinces. More precisely, in comparison to a more distant province, a geographically nearby province is assumed to be a stronger rival in competing for mobile labour and physical capital.

Apart from the variables described above, the present study also examines the impact of physical capital and human capital. Specifically, the growth of physical capital investment, ΔINV, is introduced into the growth regression under the assumption that this variable would capture an increase in technology investment and/or an increase in the number of machines used in the production process. With respect to the proxy of human capital, previous studies use such variables as the average schooling years and the enrolment rate as proxies for human capital. Due to the lack of data, in the present study, the change of human capital supply is examined instead, proxied by the change of the number of high-school graduates, $\Delta SECGRA$, and, alternatively, the change of the number of high-school and higher-education graduates, ΔGRA.

3.3.5 Data

The panel dataset contains 61 provinces for 8 years. Due to the inclusion of the lagged dependent variable and the application of the GMM estimator, the sample period reduces to 6 years and, therefore, the total number of observations is 366. Descriptive statistics and correlation matrix of the variables are presented in Tables 3.1 and 3.2. With respect to the sources of data, the road variable is calculated by using the data on road length provided by the Ministry of Transport and the data on area provided by the General Statistics Office of Vietnam; the data on the number of graduates are provided by the Ministry of Education and

Training; the remaining variables are calculated by using the relevant data provided by the General Statistics Office of Vietnam.

Table 3.1: Descriptive Statistics of Variables

Variable	Obs.	Mean	Std. Dev.	Min	Max
ΔLog(*GDPPC*)	427	0.096	0.037	-0.246	0.338
Log(*GDPPC*)	488	1.338	0.534	0.298	3.772
ΔLog(*SECGRA*)	427	0.066	0.201	-1.450	1.347
ΔLog(*GRA*)	427	0.070	0.191	-1.346	1.231
ΔLog(*INV*)	427	0.159	0.207	-0.722	1.827
Log(*ROAD*)	488	6.696	0.767	5.252	8.062
Log($w_{cont}ROAD$)	488	6.825	0.562	5.459	7.911
Log($w_{dist}ROAD$)	488	6.748	0.547	5.368	7.923
Log(*URBAN*)	488	-1.656	0.552	-2.849	-0.158
Δ*AGRGDP*	427	-0.017	0.014	-0.078	0.043
EXPGDP	488	0.298	0.448	0	3.602
IMPGDP	488	0.257	0.481	0	3.323
SOECAP	488	0.481	0.215	0.028	0.954
SOEEMP	488	0.399	0.190	0.024	0.907
SOEENT	488	0.094	0.088	0.007	0.500
ΔLog(*CPI*)	427	0.018	0.037	-0.080	0.113

Note: All monetary variables are deflated by provincial GDP deflator.

Table 3.2: Correlation Matrix of Explanatory Variables

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Log($GDPPC$)	1.0000														
2	ΔLog($SECGRA$)	-0.0425	1.0000													
3	ΔLog(GRA)	-0.0479	0.9709	1.0000												
4	ΔLog(INV)	-0.0371	0.0517	0.0543	1.0000											
5	Log($ROAD$)	0.4523	-0.0617	-0.0753	0.0093	1.0000										
6	Log($w_{cont}ROAD$)	0.3723	-0.0621	-0.0707	-0.0266	0.6805	1.0000									
7	Log($w_{dist}ROAD$)	0.3751	-0.0568	-0.0642	-0.0475	0.7164	0.9742	1.0000								
8	Log($URBAN$)	0.5864	-0.0261	-0.0239	-0.0181	0.0604	-0.1382	-0.1334	1.0000							
9	$\Delta AGRGDP$	0.1832	0.0357	0.0311	0.0213	0.0266	0.0224	0.0101	0.3050	1.0000						
10	$EXPGDP$	0.5382	-0.0238	-0.0332	-0.0062	0.1762	0.1329	0.0811	0.3868	0.1240	1.0000					
11	$IMPGDP$	0.4116	-0.0272	-0.0346	-0.0032	0.2386	0.2143	0.1546	0.3808	0.1192	0.7475	1.0000				
12	$SOECAP$	-0.4486	-0.0025	0.0019	0.0616	-0.3472	-0.4229	-0.4008	-0.0429	0.0713	-0.3762	-0.3066	1.0000			
13	$SOEEMP$	-0.3601	-0.0098	-0.0032	0.0959	-0.3017	-0.2934	-0.3059	0.0469	0.1833	-0.2893	-0.2450	0.7364	1.0000		
14	$SOEENT$	-0.5330	0.0399	0.0538	0.1174	-0.3574	-0.3475	-0.3463	-0.1962	0.0503	-0.2693	-0.2064	0.4677	0.6570	1.0000	
15	ΔLog(CPI)	-0.0164	0.3702	0.3552	-0.0011	-0.0003	0.0034	0.0068	-0.0278	0.0440	0.0011	0.0175	-0.0081	0.0356	0.0650	1.0000

Note: The sample contains 427 observations.

The road data do not provide any information on the composition of roads and, hence, cannot control for the fact that different categories of roads serve different purposes in the road network. According to the Road Traffic Act of Vietnam, the road network can be broken down into 6 categories of roads: national roads, provincial roads, district roads, urban roads, commune roads, and others. Accordingly, national roads link the Capital with other provinces' administrative centres, link three or more provincial administrative centres, and connect international ports and airports. At the local level, the road network includes provincial roads, district roads, urban roads, commune roads, and other roads. Each category of local roads is defined as follows. Provincial roads link the province's administrative centre to its district-level administrative centres and to the neighbouring provincial-level administrative centres. District roads link the administrative centre of the district to its local socio-economic centres and to the administrative centres of neighbouring districts in the province. Districts are also classified into urban, sub-urban, and rural. Urban roads are the roads within the administrative boundaries of urban districts in the province. Commune roads link the administrative centre of the commune to its villages, hamlets and equivalent units, and the administrative centres of neighbouring communes in the province.

Also, the road spillover variable used in the present study does not contain information on the composition of roads installed in the neighbouring provinces as well as the connectivity between provinces sharing administrative borders. More precisely, in regard to the connectivity, the road data are unable to illustrate whether a particular province is well-connected with the national road network, and whether provinces sharing administrative borders are well-connected with each other. In this context, it is worth recalling that, as noted previously in the literature review, the existence of a spillover effect is essentially explained by the network characteristic of infrastructure, especially transport infrastructure. What is important is the data limitations in regard to the composition of roads would affect the capacity of the road spillover variable in capturing the cross-province externalities of roads.

The use of the GMM estimator can solve the endogeneity problem associated with the infrastructure variable. The problem of endogeneity, as claimed in the related literature, may arise from the possibility that an improved endowment of infrastructure causes growth but, on the other side, growth causes a decision to improve infrastructure. However, it should be noted that in Vietnam, both provinces with growing economies and those with lagging economic performance could be targets for an increase in infrastructure provision over the period observed. In the present study, for the more developed provinces, the change in their road network which has been already denser than elsewhere was not as significant as the growth of road length in other provinces, especially in the less developed provinces.

Regarding higher-education infrastructure, the number of higher-education institutions of some provinces is time-invariant over the sample period. Furthermore, some provinces have zero values in relation to this variable, whereas others experience a growth in this category of infrastructure. Although this is understandable in practice, it is not supportive to estimate the growth impact of higher-education infrastructure using the GMM estimator. Therefore, with respect to infrastructure, the present study focuses on transport infrastructure only.

3.4 RESULTS

3.4.1 Results of the GDP Growth Model

Table 3.3 presents results of the growth impact of infrastructure, among other factors, on GDP growth at the province level. In all regressions, the null hypothesis of the first-order serial correlation test is rejected while that of the second-order serial correlation test cannot be rejected, suggesting that there is first-order serial correlation but no second-order serial correlation in the residuals. The Hansen test does not reject the null hypothesis of overidentifying restrictions. These results of the Arellano-Bond autocorrelation and the Hansen overidentifying restriction tests confirm the validity of the set of instruments and hence support the model specification.

Table 3.3: Results of the GDP Growth Model

	(1)	(2)	(3)	(4)	(5)
Log(*GDPPC*) $_{i,t-1}$	-0.039**	-0.037**	-0.037**	-0.032**	-0.035***
	(0.018)	(0.017)	(0.017)	(0.015)	(0.013)
ΔLog(*SECGRA*) $_{i,t}$	0.002	0.002		0.000	0.006
	(0.017)	(0.018)		(0.018)	(0.018)
ΔLog(*GRA*) $_{i,t}$			0.005		
			(0.022)		
ΔLog(*INV*) $_{i,t}$	0.026**	0.026**	0.025**	0.029**	0.029**
	(0.013)	(0.012)	(0.012)	(0.012)	(0.011)
Log(*ROAD*) $_{i,t}$	0.020**	0.019***	0.019***	0.018***	0.015**
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
Log(*URBAN*) $_{i,t}$	0.027	0.026	0.026	0.020	0.021
	(0.017)	(0.016)	(0.017)	(0.017)	(0.013)
Δ*AGRGDP* $_{i,t}$	-1.293*	-1.322*	-1.343*	-1.115*	-1.300*
	(0.677)	(0.692)	(0.709)	(0.660)	(0.700)
EXPGDP $_{i,t}$	0.012	0.009	0.009	0.007	0.008
	(0.021)	(0.020)	(0.019)	(0.018)	(0.017)
IMPGDP $_{i,t}$	-0.004				
	(0.014)				
SOECAP $_{i,t}$	-0.015	-0.015	-0.015		
	(0.028)	(0.029)	(0.029)		
SOEEMP $_{i,t}$				-0.001	
				(0.043)	
SOEENT $_{i,t}$					-0.156**
					(0.078)
ΔLog(*CPI*) $_{i,t}$	-0.274*	-0.279*	-0.274*	-0.271*	-0.280**
	(0.146)	(0.144)	(0.141)	(0.147)	(0.123)
Constant	-0.058	-0.051	-0.054	-0.036	-0.014
	(0.072)	(0.076)	(0.081)	(0.076)	(0.072)
Arellano-Bond AR1 Test [*p*-value]	-3.50 [0.000]	-3.47 [0.001]	-3.49 [0.000]	-3.48 [0.001]	-3.51 [0.000]
Arellano-Bond AR2 Test [*p*-value]	0.22 [0.827]	0.19 [0.851]	0.22 [0.826]	0.28 [0.779]	0.23 [0.815]
Hansen Test of Overid. Res. [*p*-value]	43.72 [0.441]	43.70 [0.357]	43.48 [0.366]	42.57 [0.403]	42.90 [0.390]
Observations	366	366	366	366	366
No. of Provinces	61	61	61	61	61

Note:

1. The dependent variable is ΔLog(*GDPPC*).
2. *GDPPC* is the predetermined variable; Δ*SECGRA*, Δ*GRA*, Δ*INV*, *ROAD*, *URBAN*, *EXPGDP* and *IMPGDP* are endogenous variables; the remaining explanatory variables are treated as exogenous.
3. Year dummies are included in all regressions.
4. Robust standard errors are in parentheses.
5. * p<0.100; ** p<0.050; *** p<0.010

A statistically significant and negative coefficient on the lagged level of the dependent variable, $GDPPC$, is obtained in all regressions. The coefficient on the variable is statistically significant at the 5 per cent level or the 1 per cent level, varying according to the specification of the growth regression. The magnitude of the coefficient ranges from -0.032 to -0.039, suggesting that a 1 per cent lower lagged level of GDP per capita raises the subsequent growth of GDP per capita by around 0.03 percentage points. This is generally supportive of a growth convergence. More precisely, after controlling for other explanatory variables, the lagged level of GDP per capita is found to have a negative effect on subsequent provincial GDP growth, providing evidence of conditional convergence to the steady state.

The growth impacts of the variables representing for the growth of human capital supply do not appear to be statistically significant in all regressions. However, this does not necessarily lead to the conclusion that human capital is not an important driver of economic progress. One reason for the weak result could be that the proxy for human capital is not appropriate. Recall that $\Delta SECGRA$ represents a change in the number of high-school graduates, and $\mathit{\Delta GRA}$ represents a change of the number of high-school and higher-education graduates. In the empirical literature, the frequently used proxies for human capital are the average schooling years or the enrolment rates. However, those data are not available for the present study.

The coefficient on the growth of physical capital investment, ΔINV, is positive and statistically significant at the 5 per cent in all regressions. The magnitude of the coefficient is at around 0.03, which indicates that an increase of 1 per cent in the growth of investment leads to an increase of around 0.03 per cent in the measure of GDP per capita growth. The finding supports the notion that the provinces with a faster growth of physical capital investment would have a greater growth performance.

Urbanisation, proxied by the proportion of urban population in total population, turns out to be of no statistical importance in explaining the growth performance. The coefficient on the variable $URBAN$ is statistically insignificant in all regressions. One explanation could be that what the variable captures might include both the economies and diseconomies associated with the concentration of people living in the urban centres, thus causing the ambiguity in the growth impact of urbanisation.

Regarding the structural transformation, it is expected that a decline in the dominance of the agriculture sector leads to a higher overall growth. In other words, a provincial economy is expected to grow faster if its structural transformation towards industrialisation proceeds faster. This is evidenced by the negative and statistically significant coefficient on the variable $\Delta AGRGDP$ obtained in all regressions. Specifically, the coefficient is statistically significant at the 10 per cent level. The magnitude of the coefficient ranges from -1.115 to -1.343, which suggests that a 1 per cent increase in the measure of this variable causes a decrease in the measure of overall growth by around 1 per cent.

With respect to the growth impact of the structural transformation in terms of ownership, the coefficients on the variables $SOECAP$ and $SOEEMP$, which represent the proportion of SOEs in physical capital stock and employment, respectively, are negative but statistically insignificant. Recall that the GDP value used in the present study also includes the proportion contributed by the state sector, and that during the period observed, the state sector still maintained a significant proportion in GDP and experienced a positive GDP growth rate as well. This might affect the results for these variables. However, there is evidence of the negative growth impact caused by an increase in the proportion of the state sector in the total number of enterprises. Specifically, the coefficient on the variable $SOEENT$ has a negative sign and turns out to be statistically significant at the 5 per cent level. The magnitude of the coefficient, -0.156, indicates that a decrease in the proportion of SOEs in total number of enterprises by 1 per cent is associated with an increase in the measure of GDP per capita growth by 0.156 per cent. This finding is in line with the fact that the process of economic reforms had led to the explosion of the private sector in terms of the number of enterprises during the sample period when the economy also achieved a high GDP growth.

There is no evidence suggesting either a positive or negative relationship between external trade and economic growth. The coefficients on all variables representing external trade, i.e. the ratio of exports to GDP, $EXPGDP$, and the ratio of imports to GDP, $IMPGDP$, are statistically insignificant in all regressions. This implies that external trade could have been a less important determinant of growth at the province level. At least in the context of Vietnam, a sectoral decomposition of exports and imports could provide an insight into the growth impact of external trade. This is because, for example, primary and non-primary exports would have different impacts on economic growth. However, the province-level data on the composition of exports and imports were not available for the present study.

Inflation, proxied by a change in the logarithm of the consumer prices index, is found to be negatively associated with the growth performance of the provincial economy. This result is supportive to the notion that macroeconomic stability, which is in part illustrated by a low inflation rate, is important for promoting economic growth. The coefficient on the variable $INFLATION$ is negative and statistically significant in all regressions. The magnitude of the coefficient ranges from -0.271 to -0.280, and the modest magnitude suggests that a 1 per cent increase in the measure of inflation leads to a decrease of 0.271 per cent in the measure of GDP per capita growth.

Of particular importance to the present study is the growth impact of transport infrastructure. All regressions show a positive relationship between the density of roads and the growth of per capita GDP. The coefficient on the variable ROAD is statistically significant at the 5 per cent level or the 1 per cent level, varying according to the model specification. The magnitude of the coefficient ranges from 0.015 to 0.020, which indicates that an increase of 1 per cent in the measure of road density results in an increase of around 0.02 per cent in the measure of GDP per capita growth.

Comparing the magnitude of the coefficients on all explanatory variables included in each regression suggests that the growth impact of roads is smaller than those of several growth determinants mentioned above. It can be that the quality, rather than the total length of roads per 1,000 km^2 of the mainland area, has a larger contribution to economic growth. However, it is not possible for the present study to explicitly control for the quality of infrastructure due to lack of the relevant data. In addition, the road variable used in the present study can capture the road network density but does not provide any information on the composition of roads and therefore cannot control for the fact that different categories of roads which serve different purposes in the road network might have different growth impacts. In particular, national roads which connect the major urban centres and result in a national road network for the whole country might have an important spatial implication for the growth performance of provincial economies since they have an important role to play, for example, in transporting finished goods to ports for export and distributing products around the country. In this way, the growth impact of national roads might differ from the impacts of other types of roads on growth.

In Table 3.4, the variable $ROAD$ is lagged for one year to minimise the endogeneity problem potentially arising from the relationship between infrastructure and economic growth, as suggested in the related literature, as well as to control for the fact that the growth impact of road construction might not happen immediately while the specification of other variables remains unchanged compared with those reported in Table 3.3. Overall, the performance of all variables is unlikely to be sensitive to this slight modification. In regard to the growth impact of roads, the coefficient on the lagged value of $ROAD$ is statistically significant at the 5 per cent level or the 1 per cent level, varying according to the specification of the growth regression. The magnitude of the coefficient ranges from 0.015 to 0.019, which is relatively similar to the results reported for the contemporaneous value of $ROAD$. This could be explained by the small variations of road length of the majority of provinces over the sample period.

Table 3.4: Results of the GDP Growth Model with the Lagged *ROAD*

	(1)	(2)	(3)	(4)	(5)
Log(*GDPPC*) $_{i,t-1}$	-0.037**	-0.037**	-0.037**	-0.032**	-0.034***
	(0.018)	(0.016)	(0.017)	(0.015)	(0.013)
ΔLog(*SECGRA*) $_{i,t}$	0.003	0.003		0.001	0.007
	(0.018)	(0.018)		(0.019)	(0.018)
ΔLog(*GRA*) $_{i,t}$			0.006		
			(0.023)		
ΔLog(*INV*) $_{i,t}$	0.026**	0.026**	0.026**	0.030**	0.030**
	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)
Log(*ROAD*) $_{i,t-1}$	0.019**	0.018***	0.018***	0.017**	0.015**
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
Log(*URBAN*) $_{i,t}$	0.025	0.025	0.025	0.020	0.020
	(0.016)	(0.016)	(0.016)	(0.016)	(0.013)
Δ*AGRGDP* $_{i,t}$	-1.269*	-1.288*	-1.307*	-1.078*	-1.276*
	(0.666)	(0.674)	(0.690)	(0.637)	(0.681)
EXPGDP $_{i,t}$	0.009	0.008	0.008	0.006	0.008
	(0.020)	(0.019)	(0.019)	(0.017)	(0.016)
IMPGDP $_{i,t}$	-0.002				
	(0.014)				
SOECAP $_{i,t}$	-0.016	-0.016	-0.016		
	(0.029)	(0.029)	(0.029)		
SOEEMP $_{i,t}$				-0.002	
				(0.042)	
SOEENT $_{i,t}$					-0.160**
					(0.076)
ΔLog(*CPI*) $_{i,t}$	-0.262*	-0.266**	-0.260**	-0.257*	-0.272**
	(0.134)	(0.134)	(0.132)	(0.137)	(0.118)
Constant	-0.048	-0.045	-0.048	-0.031	-0.011
	(0.069)	(0.073)	(0.078)	(0.073)	(0.069)
Arellano-Bond AR1 Test [*p*-value]	-3.50 [0.000]	-3.47 [0.001]	-3.50 [0.000]	-3.47 [0.001]	-3.51 [0.000]
Arellano-Bond AR2 Test [*p*-value]	0.28 [0.781]	0.25 [0.800]	0.29 [0.775]	0.37 [0.714]	0.29 [0.775]
Hansen Test of Overid. Res. [*p*-value]	44.33 [0.584]	45.86 [0.436]	46.16 [0.424]	45.21 [0.463]	48.36 [0.339]
Observations	366	366	366	366	366
No. of Provinces	61	61	61	61	61

Note:

1. The dependent variable is ΔLog($GDPPC$).
2. $GDPPC$ is the predetermined variable; $ROAD$ is treated as predetermined; $\Delta SECGRA$, ΔGRA, ΔINV, $URBAN$, $EXPGDP$ and $IMPGDP$ are endogenous variables; the remaining explanatory variables are treated as exogenous.
3. Year dummies are included in all regressions.
4. Robust standard errors are in parentheses.
5. * p<0.100; ** p<0.050; *** p<0.010

3.4.2 Results of the GDP Growth Model with Road Spillover Variable

Table 3.5 presents results for the estimation of the model including the average road density in the neighbouring provinces as an additional explanatory variable. As stated above, the road spillover variable is supposed to capture the spillover effect of roads installed in the surrounding area of a particular province. Two alternative spatial weights matrices have been tried to measure the variable $wROAD$: the contiguity-based spatial weights matrix, w_{cont}, and the contiguity- and distance-based spatial weights matrix, w_{dist}. The former is adopted to measure the road spillover variable reported in regressions (1)-(5), while the latter is in regressions (6)-(10). In all regressions, the system GMM estimation shows that there is no evidence of second order serial correlation in the first-differenced residuals and the Hansen test does not reject the validity of instruments, all of which results suggest the consistency of the estimator being used.

Table 3.5: Results of the GDP Growth Model with Road Spillover Variable

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(*GDPPC*) $_{i,t-1}$	-0.039**	-0.037**	-0.037**	-0.034**	-0.035***	-0.035*	-0.035**	-0.035**	-0.031*	-0.033**
	(0.018)	(0.017)	(0.017)	(0.016)	(0.013)	(0.018)	(0.017)	(0.017)	(0.016)	(0.014)
ΔLog(*SECGRA*) $_{i,t}$	0.003	0.004		0.002	0.007	0.004	0.004		0.002	0.007
	(0.017)	(0.017)		(0.018)	(0.017)	(0.017)	(0.017)		(0.018)	(0.017)
ΔLog(*GRA*) $_{i,t}$			0.007					0.008		
			(0.021)					(0.021)		
ΔLog(*INV*) $_{i,t}$	0.025**	0.026**	0.025**	0.031**	0.029**	0.023*	0.024*	0.023*	0.030**	0.027**
	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)
Log(*ROAD*) $_{i,t}$	0.022**	0.021**	0.021**	0.018**	0.018**	0.025**	0.025**	0.025**	0.022**	0.021**
	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.012)	(0.011)	(0.011)	(0.010)	(0.011)
Log(w_{cont}*ROAD*) $_{i,t}$	-0.003	-0.004	-0.004	-0.000	-0.004					
	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)					
Log(w_{dist}*ROAD*) $_{i,t}$						-0.011	-0.011	-0.011	-0.007	-0.010
						(0.012)	(0.012)	(0.012)	(0.011)	(0.011)
Log(*URBAN*) $_{i,t}$	0.026	0.025	0.025	0.023	0.020	0.022	0.022	0.022	0.019	0.017
	(0.017)	(0.017)	(0.018)	(0.018)	(0.014)	(0.017)	(0.017)	(0.018)	(0.018)	(0.014)
Δ*AGRGDP* $_{i,t}$	-1.305*	-1.329*	-1.351*	-1.146*	-1.312*	-1.285*	-1.306*	-1.328*	-1.107*	-1.288*
	(0.675)	(0.700)	(0.715)	(0.677)	(0.704)	(0.661)	(0.682)	(0.697)	(0.658)	(0.687)
EXPGDP $_{i,t}$	0.012	0.009	0.009	0.007	0.009	0.009	0.008	0.008	0.007	0.008
	(0.021)	(0.020)	(0.020)	(0.018)	(0.018)	(0.021)	(0.020)	(0.020)	(0.018)	(0.017)
IMPGDP $_{i,t}$	-0.003					-0.001				
	(0.014)					(0.014)				
SOECAP $_{i,t}$	-0.016	-0.016	-0.016			-0.017	-0.017	-0.017		
	(0.029)	(0.030)	(0.030)			(0.029)	(0.029)	(0.029)		
SOEEMP $_{i,t}$				-0.002					-0.001	
				(0.043)					(0.043)	
SOEENT $_{i,t}$					-0.154*					-0.156*
					(0.079)					(0.080)
ΔLog(*CPI*) $_{i,t}$	-0.277*	-0.281*	-0.276*	-0.273*	-0.282**	-0.277*	-0.282*	-0.278*	-0.275*	-0.284**
	(0.147)	(0.145)	(0.142)	(0.148)	(0.123)	(0.147)	(0.145)	(0.142)	(0.148)	(0.123)

Constant	-0.047	-0.040	-0.044	-0.043	-0.006	-0.013	-0.012	-0.016	-0.013	0.019
	(0.084)	(0.092)	(0.097)	(0.090)	(0.084)	(0.084)	(0.091)	(0.096)	(0.091)	(0.083)
Arellano-Bond AR1 Test [p-value]	-3.46 [0.000]	-3.47 [0.001]	-3.47 [0.001]	-3.49 [0.000]	-3.52 [0.000]	-3.51 [0.000]	-3.47 [0.001]	-3.49 [0.000]	-3.49 [0.000]	-3.52 [0.000]
Arellano-Bond AR2 Test [p-value]	0.19 [0.847]	0.19 [0.851]	0.23 [0.820]	0.27 [0.786]	0.24 [0.814]	0.24 [0.807]	0.22 [0.827]	0.26 [0.820]	0.30 [0.768]	0.26 [0.799]
Hansen Test of Overid. Res. [p-value]	42.31 [0.414]	43.70 [0.357]	41.72 [0.439]	41.82 [0.435]	39.96 [0.517]	41.19 [0.550]	41.33 [0.456]	41.53 [0.448]	41.74 [0.438]	39.59 [0.533]
Observations	366	366	366	366	366	366	366	366	366	366
No. of Provinces	61	61	61	61	61	61	61	61	61	61

Note:

1. The dependent variable is $\Delta \text{Log}(GDPPC)$.
2. $GDPPC$ is the predetermined variable; $\Delta SECGRA$, ΔGRA, ΔINV, $ROAD$, $URBAN$, $EXPGDP$ and $IMPGDP$ are endogenous variables; the remaining explanatory variables are treated as exogenous.
3. Year dummies are included in all regressions.
4. Robust standard errors are in parentheses.
5. * p<0.100; ** p<0.050; *** p<0.010

The inclusion of the road spillover variable leads to a small change in the magnitude of the variable $ROAD$, compared with that reported previously in Table 3.3. However, there is no evidence of the cross-province spillover effect of roads. One explanation could be based on the high correlation between the variables $ROAD$ and $wROAD$ presented in Table 3.2. On the other side, while the literature claims that the spillovers of transport infrastructure across geographical areas are explained by its network characteristic, the road data used in the present study do not contain information on the connectivity between provinces sharing administrative borders as well as the connectivity of the provincial road network into the national road network.

Another reason for the existence of the spillover effect of infrastructure across areas is that differences in infrastructure endowments across areas could trigger regional competition for production factors, as presented in the literature review. Previous studies find empirical evidence on the negative spatial externalities of infrastructure under the hypothesis that regions with improved infrastructure services draw production factors and economic activities away from their neighbours. In Vietnam, during the sample period the growth rate of road length was, on average, higher in less-developed regions, which can be interpreted as an attempt to improve the infrastructure endowment of these provinces as well as to connect them into the national road network. An increased infrastructure endowment alone might be not enough to generate a strong negative spillover effect since the literature emphasises that infrastructure endowment is necessary but not sufficient to promote economic development and enhance competitiveness of a particular area.

Furthermore, as Boarnet (1998) argues, the negative spillover effect of transport infrastructure on growth also depends on whether the adjacent areas can perform as a substitute for a particular area in terms of a production base. This in turn implies the importance of controlling for the similarity in specialisation structure of a particular province and its neighbouring provinces. However, the construction of the road spillover variable using the GDP-composition-based spatial weights matrix, which can capture the specialisation structure, may be problematic on econometric grounds. As presented in the literature review, the matrix used to construct the spillover variable must not contain any of the exogenous or endogenous variables used in the regression to ensure that the empirical model not become highly non-linear (e.g., Anselin and Bera, 1998; Abreu *et al.*, 2005).

From the geographical perspective, the S-shaped mainland of Vietnam might imply another difficulty for studying the cross-province spillover effects of the road network. In addition, the number of neighbouring provinces varies significantly across provinces, ranging from 2 to 7. Also, some provinces share common boundaries with provinces that are located in the same geographical region while others share common boundaries with both provinces locating in the same geographical region and provinces located in the neighbouring region. Recall that, as presented in Chapter 2, there are significant geographical disparities in a wide range of aspects, especially infrastructure endowment and economic development at both province and region levels. Additionally, there are also differences in terms of terrain particularly across regions which are influential factors determining the specific features of the road network installed in each region. In this context, the spatial weights matrices used to calculate the average density of road network installed in neighbouring provinces might not fully capture the diversity and the complexity in regard to the neighbour relations between Vietnamese provinces.

In Table 3.6, the road variable and the road spillover variable are lagged for one year to control for the fact that the growth impact of road construction might not happen immediately while the specification of other variables remains unchanged compared with those reported in Table 3.5. This is, in addition, aimed at minimising the problem of endogeneity which, as the literature suggests, can arise from the relationship between infrastructure installed in a particular province and economic growth in that province. Overall, the performance of all variables is unlikely to be sensitive to this slight modification of the growth regressions. In regard to the growth impacts of roads and neighbouring roads, the magnitude of the coefficients on the lagged variables is relatively similar to the results reported in Table 3.5 for the contemporaneous variables. This again could be explained by the small variations of road length of the majority of provinces over the sample period.

Table 3.6: Results of the Model of GDP Growth with the Lagged *ROAD* and the Lagged *wROAD*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(*GDPPC*) $_{i,t-1}$	-0.038**	-0.037**	-0.037**	-0.034**	-0.035***	-0.034*	-0.035**	-0.035**	-0.031*	-0.033**
	(0.018)	(0.017)	(0.017)	(0.015)	(0.013)	(0.018)	(0.017)	(0.017)	(0.016)	(0.013)
ΔLog(*SECGRA*) $_{i,t}$	0.004	0.004		0.003	0.008	0.005	0.005		0.003	0.008
	(0.018)	(0.018)		(0.019)	(0.018)	(0.018)	(0.018)		(0.018)	(0.018)
ΔLog(*GRA*) $_{i,t}$			0.007					0.008		
			(0.022)					(0.022)		
ΔLog(*INV*) $_{i,t}$	0.026**	0.027**	0.026**	0.032**	0.030**	0.025*	0.025**	0.024*	0.031**	0.028**
	(0.013)	(0.012)	(0.013)	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)
Log(*ROAD*) $_{i,t-1}$	0.020**	0.020**	0.020**	0.017**	0.017*	0.023**	0.023**	0.023**	0.020**	0.020*
	(0.010)	(0.009)	(0.009)	(0.008)	(0.009)	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)
Log(w_{cont}*ROAD*) $_{i,t-1}$	-0.003	-0.003	-0.003	0.000	-0.003					
	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)					
Log(w_{dist}*ROAD*) $_{i,t-1}$						-0.009	-0.009	-0.009	-0.006	-0.009
						(0.011)	(0.012)	(0.012)	(0.011)	(0.011)
Log(*URBAN*) $_{i,t}$	0.025	0.025	0.025	0.022	0.020	0.021	0.022	0.022	0.018	0.017
	(0.016)	(0.017)	(0.017)	(0.018)	(0.014)	(0.017)	(0.017)	(0.017)	(0.018)	(0.014)
Δ*AGRGDP* $_{i,t}$	-1.286*	-1.301*	-1.321*	-1.118*	-1.291*	-1.262*	-1.273*	-1.293*	-1.075*	-1.263*
	(0.664)	(0.682)	(0.698)	(0.656)	(0.685)	(0.652)	(0.665)	(0.680)	(0.637)	(0.669)
EXPGDP $_{i,t}$	0.010	0.008	0.008	0.007	0.008	0.007	0.007	0.007	0.006	0.007
	(0.021)	(0.019)	(0.019)	(0.018)	(0.017)	(0.020)	(0.019)	(0.019)	(0.018)	(0.017)
IMPGDP $_{i,t}$	-0.001					0.001				
	(0.014)					(0.014)				
SOECAP $_{i,t}$	-0.016	-0.016	-0.016			-0.017	-0.018	-0.018		
	(0.029)	(0.030)	(0.030)			(0.029)	(0.029)	(0.029)		
SOEEMP $_{i,t}$				-0.003					-0.002	
				(0.042)					(0.042)	
SOEENT $_{i,t}$					-0.159**					-0.161**
					(0.077)					(0.078)
ΔLog(*CPI*) $_{i,t}$	-0.262*	-0.266**	-0.261**	-0.255*	-0.272**	-0.263*	-0.266**	-0.261**	-0.257*	-0.274**
	(0.135)	(0.134)	(0.132)	(0.137)	(0.118)	(0.136)	(0.134)	(0.133)	(0.138)	(0.118)

Constant	-0.042	-0.039	-0.043	-0.042	-0.004	-0.010	-0.012	-0.016	-0.013	0.021
	(0.079)	(0.088)	(0.093)	(0.087)	(0.080)	(0.079)	(0.088)	(0.092)	(0.088)	(0.080)
Arellano-Bond AR1 Test [p-value]	-3.49 [0.000]	-3.46 [0.001]	-3.48 [0.000]	-3.48 [0.000]	-3.51 [0.000]	-3.50 [0.000]	-3.47 [0.001]	-3.49 [0.000]	-3.48 [0.001]	-3.52 [0.000]
Arellano-Bond AR2 Test [p-value]	0.28 [0.781]	0.26 [0.797]	0.29 [0.820]	0.36 [0.722]	0.29 [0.771]	0.31 [0.760]	0.29 [0.773]	0.26 [0.820]	0.39 [0.699]	0.32 [0.751]
Hansen Test of Overid. Res. [p-value]	43.18 [0.631]	45.83 [0.437]	47.79 [0.439]	45.00 [0.472]	47.84 [0.357]	44.14 [0.592]	45.44 [0.453]	41.53 [0.448]	44.99 [0.472]	47.08 [0.387]
Observations	366	366	366	366	366	366	366	366	366	366
No. of Provinces	61	61	61	61	61	61	61	61	61	61

Note:
1. The dependent variable is $\Delta \text{Log}(GDPPC)$.
2. $GDPPC$ is the predetermined variable; $ROAD$ is treated as predetermined; $\Delta SECGRA$, ΔGRA, ΔINV, $URBAN$, $EXPGDP$ and $IMPGDP$ are endogenous variables; the remaining explanatory variables are treated as exogenous.
3. Year dummies are included in all regressions.
4. Robust standard errors are in parentheses.
5. * p<0.100; ** p<0.050; *** p<0.010

3.5 SUMMARY AND POLICY DISCUSSION

The empirical findings obtained in Chapter 3 is summarised below.

First of all, the empirical results lead to the conclusion that transport infrastructure has an important role to play in promoting growth. The possibility that the provision of infrastructure in one province also affects economic performance of neighbouring provinces has been explicitly taken into account in the empirical analysis. Alternative spatial weights matrices have been tried to measure the road spillover variable. However, there is no strong evidence supporting the hypothesis on the existence of the cross-province spillover effect of transport infrastructure. Explanations can be based on the nature of the road data used in the present study, the growth patterns of roads and GDP over the sample period, and the geographical characteristics of Vietnamese provinces.

Also, the empirical analysis uses a number of other explanatory variables to control for the differences between provincial economies in regard to the performance of inflation measured by a change of the logarithm of the consumer price index, the transformation of the GDP sectoral structure, the ownership structure, the degree of urbanisation, and the external trading activities. In particular, inflation turns out to be negatively associated with growth. A faster shift towards the economy which is less dependent on agriculture leads to a higher overall growth; whereas, there is no evidence indicating that urbanisation is a factor behind growth and this ambiguity in regard to the growth impact of urbanisation could be due to the fact that the variable captures might include both economies and diseconomies associated with the concentration of people living in the urban centres. Variables proxying for external trade play no role in explaining growth performance, and this might suggest that the composition of exports and imports could be more important for growth. Regarding the growth impacts of variables representing the ownership structure of the provincial economies, the empirical results show that a decrease in the proportion of SOEs is positively associated with an increase in the overall growth.

Besides, the growth of physical capital investment appears to be an important driver of economic growth. However, the relationship between the growth of human capital supply and the growth of GDP per capita is unable to be identified. Finally, to some extent, the empirical result for the lagged level of GDP per capita confirms the convergence hypothesis, that is, the poorer provincial economies grow faster than the richer ones.

Several policy implications can be interpreted from the main findings summarised above.

Firstly, policies which are aimed at increasing the density of the road network, i.e. the length of roads per 1,000 km^2 of the provincial area, to accelerate economic growth is generally supported by the empirical evidence found in this chapter. However, it is important to recall that the expectation for a positive growth impact of infrastructure is based on a number of assumptions which are essentially related to the quality of infrastructure. In particular, it is assumed that an increased road density is an indicator of the road network improvement. Then, an improvement of infrastructure quality would be conductive to economic growth. It is also important to note that the results obtained for the road variable should be interpreted on the basis of the physical presence of infrastructure, and the results do not necessarily lead to any suggestion in relation to public spending on road infrastructure.

Secondly, the negative association between the change of the proportion of agriculture in GDP and the change of per capita GDP will naturally lead to a policy implication that the structural transformation toward industrialisation has an important role to play in accelerating overall economic growth. As presented in Chapter 2, the structure of the Vietnamese economy has been shifting from agriculture-dominated to non-agriculture-dominated. Specifically, the agriculture sector has been a declining sector in terms of its decreased share in total GDP. However, it is not necessary that the empirical result on the growth-enhancing impact of the structural transformation toward industrialisation leads to an underestimation of the role of the agriculture sector in the Vietnamese economy.[28]

Thirdly, policies which are aimed at reducing the volatility of the provincial economy to macroeconomic shocks, particularly the shocks associated with the CPI performance, may be critical for achieving the growth target. This shows up in the empirical result through the negative influence of the change of the logarithm of the consumer price index on the change of GDP per capita. This also supports the widely acknowledged notion that although macroeconomic stability alone cannot promote growth, the economy cannot grow unless the macroeconomic environment is favourable or stable.

Finally, in relation to the growth implication of the ownership structure, there is some evidence supporting the notion that the restructuring towards a decline in the size of the SOE sector relative to the whole economy would be beneficial for the overall growth performance. More precisely, the empirical results suggest that the on-going SOE restructuring and the implementation of policies aimed at encouraging the private sector development would be conductive to the growth of the whole economy. The subsequent empirical chapters are

[28] This argument can be justified as follows. First of all, an extremely large proportion of the population still rely on agriculture to earn their livings, which is supposed to be illustrated by the large pools of rural population and labour as shown in Chapter 2. Besides, Vietnam has competitive advantages in producing a wide range of agricultural products which are also exportable. It is important to note that the empirical result is obtained as average for all provinces and hence ignores the cross-province variations in comparative advantages. In practice, for the provinces that have comparative advantages in agriculture, policies to promote a productive agriculture sector would also be growth-enhancing. Then, the structural transformation of provincial economies should be in accordance with the principle of comparative advantages. Last but not least, it is worth noting that, a sustainable economic growth has to depend on a balanced development of the agriculture sector and the non-agricultural sectors. This is due to the fact that a productive agriculture sector not only contributes to the development of other sectors, which need the supply of materials from agricultural industries, but also plays an essential role in ensuring national food security.

devoted to analysing two aspects of the private sector, namely, employment and location choices, which focus on the importance of infrastructure but also control for the impact of other factors including the presence of SOEs in the economy.

CHAPTER 4

INFRASTRUCTURE AND EMPLOYMENT

4.1 GENERAL BACKGROUND

Basically, the competitive labour market theory maintains that the equilibrium levels of employment and wages are determined by the interaction between the labour demand and the labour supply. According to Eberts and Stone (1992), the equilibrium levels remain unchanged unless there is a "shock" to the labour market, and public infrastructure investment is an example of that "shock". Specifically, they argue that the public provision of infrastructure could lead to an enhancement of the attractiveness of a region and thus affect the interaction between the demand for labour and the supply of labour. On the labour demand side, infrastructure could be viewed as an unpaid factor of production as well as a productivity-enhancing factor, then affecting the demand for labour. In addition, infrastructure could attract both new establishments and firms re-locating from elsewhere, thus raising the labour demand in the local labour market. On the labour supply side, local amenities enhanced by infrastructure investment could also affect the residential location decision. After all, any adjustment in the labour demand and the labour supply would lead to the adjustment of the equilibrium employment level and the equilibrium wage level.

From a spatial perspective, the creation of jobs in a particular area could be at the expense of employment in other areas where some of those jobs were previously located. It is suggested that differences in infrastructure endowment could provide an explanation for disparities in employment across geographical areas (e.g., Seitz, 2000; Cohen and Paul, 2004). Then, the overall employment impact of infrastructure can be ambiguous at the national level.

There is an extensive, but inconclusive, empirical literature on the relationship between infrastructure and employment. Specifically, a positive employment impact of infrastructure has been identified in some studies, whereas the statistically significant, or even negative, relations between infrastructure and employment have also been found in other studies. The empirical literature on the spillover effect of infrastructure across areas is growing, but the majority of previous studies have been concerned with the impact of the infrastructure spillovers on productivity and output growth.

Private sector development has been one crucial aspect of the economic reforms in Vietnam. As presented in Chapter 2, the development of the private sector was most impressive in the increased number of non-SOEs, i.e. non-state enterprises and foreign-invested enterprises, and hence the increased number of people employed by these enterprises. According to the Annual Enterprise Surveys, SOEs accounted for an average share of 41.12 per cent per year in the total number of employees during the period 2000-2007. However, the workforce of the SOE sector was declining annually at an average rate of -2.24 per cent. On the contrary, during the same period, the non-state and foreign-invested sectors, experienced high average rates of annual growth of 21.02 per cent and 22.73 per cent, respectively, in their number of employees. Given that the private sector has played an increasingly important role in the economy in general and in the labour market in particular, it is worth identifying determinants of employment in the private sector.

Chapter 4 examines factors determining private sector employment in Vietnamese provinces during the period 2000-2007. Among various potential determinants of employment, the focus is placed on transport infrastructure. While focusing on the employment impact of transport infrastructure, the empirical analysis also examines the existence of spillovers from transport infrastructure installed in neighbouring provinces. In addition, the specification of the empirical model also control for the fact that infrastructure is just one of a range of potential factors influencing private sector employment. Empirical results obtained in this chapter are expected to contribute to understanding not only the role of infrastructure but also the importance of the redefining of ownership structure, among other factors, for stimulating private sector employment. Regarding methodology, the Feasible Generalised Least Squares (FGLS) estimator is applied with a control for heteroskedasticity and panel-specific first-order autocorrelation.

The rest of Chapter 4 is organised as follows. Section 4.2 describes a conceptual framework to analyse the impact of transport infrastructure on labour demand and labour supply, followed by a survey of empirical findings. Section 4.3 presents the model specification and estimation methods, describes variables added in the empirical models, and discusses issues related to the data used in the empirical analysis. Results are presented and discussed in Section 4.4. Finally, Section 4.5 summarises this chapter and draws policy implications.

4.2 LITERATURE REVIEW

The literature review begins with a conceptual framework explaining how transport infrastructure affects the labour market. While the relationship between transport infrastructure and employment has been examined in an enormous number of studies, the empirical findings are mixed. Then, the review of the empirical literature focuses on the main issues learnt from previous studies, which in turn provide some suggestions for further research.

4.2.1 The Conceptual Framework

4.2.1.1 Basic Principle of the Labour Market

The basic principle of the competitive labour market theory states that, in a local labour market, the equilibrium levels of employment and wages are determined by the interaction between labour demand and labour supply. The following mathematical description of the wage and employment equilibrium levels is drawn from Benjamin *et al.* (2005). It is assumed that the labour market is competitive. In addition, firms are assumed to be able to obtain all the labour they wish at the going wage rate and the market labour demand curve is defined as the horizontal summation of the labour demand curves of the individual firms. Also, the market labour supply curve is assumed to be the horizontal summation of the labour supply curves of the individuals. Then, the labour demand and supply curves can be mathematically expressed as, respectively,

$$LD = \alpha + \beta W + \gamma X \qquad (4.1)$$

and

$$LS = \varphi + \omega W + \mu Z \tag{4.2}$$

where LD and LS are the labour demand and the labour supply, respectively; W is the wage rate; X and Z represent factors determining the labour demand and the labour supply, respectively.

Solving the structural labour demand and labour supply equations gives the reduced-form equations for the equilibrium wage level and the equilibrium employment level. In particular, equating the labour demand equation and the labour supply equation, which are specified above, produces the following equation

$$\alpha + \beta W + \gamma X = \varphi + \omega W + \mu Z \tag{4.3}$$

which can be re-arranged to produce the reduced-form equation for wages

$$W = \frac{(\alpha + \gamma X - \varphi - \mu Z)}{(\omega - \beta)} \tag{4.4}$$

Then, substituting W into $L = LD = \alpha + \beta W + \gamma X$ to get

$$L = LD = \alpha + \gamma X + \beta \left[\frac{(\alpha + \gamma X - \varphi - \mu Z)}{(\omega - \beta)} \right] \tag{4.5}$$

which can be re-arranged to obtain the reduced-form equation for employment

$$L = \frac{(\alpha\omega + \gamma\omega X - \beta\varphi - \beta\mu Z)}{(\omega - \beta)} \quad (4.6)$$

where L represents employment.

The wage and employment effects of a change in X and/or a change in Z can be determined from these reduced-form equations. According to Benjamin et al. (2005), an increase in X by 1 unit causes the equilibrium wage level to increase by $\gamma/(\omega - \beta)$ and the equilibrium employment level to increase by $\gamma\omega/(\omega - \beta)$. Since the parameter β is assumed to have a negative sign, that is, wages are assumed to have a negative effect on the labour demand, the divisor $(\omega - \beta)$ is positive.

According to Eberts and Stone (1992), the equilibrium levels of employment and wages in a regional labour market remains unchanged unless the regional labour market is disturbed by a shock which may stem from either inside or outside a region. They define "a shock" as "an event" that "alters the steady-state growth path of each regional labour market" (Eberts and Stone, 1992, p.47). The responses of the regional labour markers to those "shocks", which are also called "market disturbances" (Eberts and Stone, 1992, p.14), are divided into two groups, that is, the responses of households and the responses of firms. The former are divided further into three categories, namely, employment changes resulting from shifts in unemployment, labour force participation changes, and population changes due to migration. The latter include changes in employment resulting from openings, expansions, contractions and closures. Examples of exogenous shocks, which can affect more than one region simultaneously, although whose the net result might differ by region, include increases in oil prices and foreign competition, changes in the age structure and other characteristics of the regional labour pool, innovations in industry-wide technology, and fluctuations in the aggregate demand for the products of the region. On the other side, examples of internal shocks, which are supposed to be specific to the region, include new technology development and implementation specifically within the region, improved labour quality gained through training programmes, and improvement in amenities installed in the region such as infrastructure.

With respect to the role of infrastructure as a shock to the regional labour market, Eberts and Stone (1992) argue that the investment in public capital stock, such as roads, highways and water treatment facilities, would offer firms located in a particular area a competitive advantage over firms in other areas which do not provide the same public services at comparable tax costs. Furthermore, infrastructure investment can lead to adjustments in the demand for labour and/or the supply of labour in a particular area, leading to changes in the equilibrium levels of local employment and wages. How transport infrastructure improvement affects the labour demand and supply is discussed further in the next section.

Also according to Eberts and Stone (1992), one relevant issue for understanding the process of the market labour adjustment is the variability between wages and employment. Holzer and Montgomery (1990), for example, argue that firms are more likely to change their employment rather than the wages payable to their employees in response to the shocks to the local labour market. Moreover, as Eberts and Stone (1992) note, if wages are less sensitive than employment to changes in the labour demand, then the labour market adjustment may essentially come through employment. Therefore, the present study focuses on the employment impact of infrastructure.

4.2.1.2 Transport Infrastructure and Labour Demand

On the relationship between transport infrastructure and the production process, the basic premise is that an improvement in infrastructure can raise output and productivity in both direct and indirect ways. Directly, infrastructure is viewed as an intermediate input that enters into a firm's production process in the same way as other inputs, such as labour and physical capital. Indirectly, the impact of infrastructure on the production process arises from the possibility that infrastructure can be supportive to the productivity of other inputs used by the firm which in turn contributes to the overall productivity. In this sense, infrastructure is complementary to private inputs. After all, any changes in productivity will generate new production possibilities for the firm, leading to adjustments in inputs, including labour, used to produce outputs at the market demand.

Rietveld and Bruinsma (1998) show that, the impact of transport infrastructure improvement on employment may occur via not only the complementary, but also the substitution, relations between labour and other production inputs, such as physical capital. Concerning the substitution effect, an enhancement in productivity, contributed by infrastructure improvement, may encourage a shift from a labour-intensive to a capital-intensive (or, from an unskilled-intensive to a skilled-intensive) production process. Therefore, an increase in productivity may lead to a decrease in employment since a smaller amount of labour input is required to produce a given level of output demanded. Put differently, the labour demand of the firm may decline if the market demand of its products remain unchanged or does not increase while its productivity is enhanced in part thanks to the productive effect of transport infrastructure improvement. In this way, infrastructure improvement is a substitute for labour. Due to the uncertainty of the relative extent of this substitution effect, the net effect of infrastructure improvement on employment may be ambiguous.

An improvement in transport infrastructure can also have competition effects and spatial implications that entail the uncertainty in predicting the net effect on labour demand in a particular area. With respect to competition, wider market access associated with transport infrastructure improvement can facilitate increased interactions between firms and hence generate increased competition, potentially leading to further increased productivity. Market share will then be redistributed to the advantage of those firms which are more productive and hence able to adapt to the expanded market. This can consequently affect the labour demand.

In a broader dimension, the net effect of interregional competition facilitated by transport infrastructure improvement on local labour demand is even more ambiguous. Job losses can happen due to an increase in interregional trading activities which in turn may lead to the closure of less competitive firms, and even a less competitive sector, previously located in a particular area. As in the case of intra-regional competition but in a broader geographical dimension in which the market is no longer limited by jurisdictional boundaries, there would be changes in the distribution of market share across firms and hence their labour demand.

The locational behaviour of firms is another source explaining changes in labour demand. Theories and empirical studies on firm location choices point to the importance of two broad sets of factors shaping the choices of firms to locate in a particular area, that is, the size of the local market and the production costs, emphasising the important role of transport infrastructure. Specifically, transport infrastructure improvement can lead to a reduction of transport costs which, according to the classical literature of industrial location, play a significant role in determining the choices of firms seeking to locate where the costs for circulating inputs and outputs from factories to markets, and vice versa, are minimised. On the other hand, transport improvement implies a better access to a particular area and, thus, its market. Of particular interest here are the implications of the location choices of firms for job creation. A higher productivity of business environment, contributed by transport infrastructure improvement, in a particular area can attract the entry of new establishments and the relocation of firms from elsewhere, resulting in an increase in overall production and hence a higher labour demand. However, in the case of relocation, the creation of jobs in one area could be at the expense of job creation in other areas where some of those jobs were previously located. Then, transport infrastructure improvement may not lead to an overall increase in the labour demand at the national level.

4.2.1.3 Transport Infrastructure and Labour Supply

In addition to its impact on productivity, intra-regional and inter-regional competition, as well as the locational behaviour of firms, which could lead to changes in the labour demand, transport infrastructure improvement can affect the supply of labour through its impact on labour-market accessibility and inter-regional labour mobility. On the one side, the reduced travel time and costs associated with transport infrastructure improvement, (e.g., Bruinsma, 1995; Buurman and Rietveld, 1999), can contribute to enhancing the accessibility to job opportunities, then encouraging people to participate in the labour force. On the other side, transport costs can also play an important role in shaping residential location choices as well as decisions of households to migrate.[29]

At the sub-national level, the labour supply will not be limited geographically. In particular, improvement of accessibility associated with improvement in transport infrastructure will allow job seekers to travel across jurisdictional boundaries. Furthermore, a reduction in

[29] According to the residential location literature, a household's choice of residence is influenced by the household's individual characteristics (e.g., Gabriel and Rosenthal, 1989) and the attributes, particularly public services, of the area which the household chooses to reside (e.g., Freidman, 1981; Quigley, 1985; Nechyba and Strauss, 1998).

commuting costs would enable workers to increase the area of job search and make longer journeys for a given amount of generalised costs (e.g., Rietveld, 1994; Vickerman, 2002). In addition, since people can be recruited to work in a particular area without changing their residence status, the labour pool of neighbouring regions may be an additional source of labour supply and this can be explained by the proximity between regional labour markets.

In the long run, transport infrastructure improvement can attract in-migration, which in turn contributes to an increased size of local population as well as an increased number of people available to participate into the local labour force. Given that there is no restriction on labour mobility, especially at the sub-national level, the in-migration in response to transport infrastructure improvement will bring both skilled and unskilled job seekers and thus intensify the competition for jobs in the local market.

As a result of these impacts of transport infrastructure improvement on an increase in the supply side of the local labour market, an enhancement in labour productivity is also probably attributable to a better match between the demand for and the supply of skilled workers. In addition, transport infrastructure improvement affects not just labour as a production input, but also as a factor constituting the final demand and, therefore, leads to an adjustment of the market size. More precisely, firms located in the areas with an increased in-migration will have more choices in terms of employment and, at the same time, will have opportunities to supply the market in part expanded by the entry of people moving from elsewhere. In this way, transport infrastructure improvement not only encourages the expansion of existing firms but also attracts the location choices of new firms indirectly by facilitating in-migration.

Nevertheless, the locational advantage of a particular area could decrease as a consequence of the provision of transport infrastructure. For instance, improvement in transport infrastructure may induce more traffic, leading to congestion and then bidding up business costs, beyond other environment-related adverse effects. In addition, transport infrastructure improvement may raise the values of land and property, which in turn may represent a cost barrier that not only limits the movement of households and firms into those areas but also discourages the expansion of existing firms (Cameron and Muellbauer, 1998).

Related questions are as to whether jobs follow people or people follow jobs, or whether the location choices of firms and households are simultaneous. Firms may prefer to locate near the markets for their inputs, including labour, and outputs, and this implies that jobs follow people. Given that there is a link between the concentration of firms and job opportunities, improvement in transport infrastructure can improve job accessibility for local residents as well as encourage people living in elsewhere to move in a particular area having such the improvement to seek for jobs. In this case, people follow jobs. Besides, this migration response could lead to an increase in population, which represents not only an enlarged pool of potential customers but also an increased number of potential job seekers in the local market. As noted previously, this in turn attracts the location choices of firms and then increases the labour demand in the local market.

Last but not least, due to its network characteristics, transport infrastructure can also generate the effects that spill over across geographical areas, affecting both the labour demand and the labour supply in adjacent or more distant areas. This implies that the employment impact is not necessarily distributed evenly over space. On the demand side, since transport infrastructure improvement can alter the competitive advantages across geographical areas, leading to the spatial redistribution of economic activities and then job opportunities. On the supply side, improvement in transport infrastructure will allow job seekers to travel across jurisdictional boundaries as well as influence the residential location choices and migration decisions of households and, therefore, the labour supply will not be limited geographically.

4.2.2 Empirical Literature on the Employment Impact of Transport Infrastructure

The impact of transport infrastructure on employment has been investigated in a number of empirical works using different levels of data, different approaches of empirical modelling, and a variety of econometric techniques. Since transport infrastructure is just one of a number of potential factors contributing to changes in the location of employment, these empirical works statistically isolate the effect of transport infrastructure while controlling for the effects caused by other factors. The review of the empirical literature presented in this section summarises the main issues learnt from previous studies, which in turn provide some suggestions for further research.

Firstly, empirical studies can be differentiated according to the proxies of infrastructure employed. The majority of studies use either the monetary or physical measures, or both.[30] Many studies examine the employment impact of government infrastructure expenditure to consider whether there is empirical evidence supporting the justification of job creation for public spending decisions. Some of these studies measure the stock of transport infrastructure in monetary terms (e.g., Nadiri and Mamuneas, 1998; Dalenberg *et al.*, 1998; and Haughwout, 1999). Regarding the physical measures of transport infrastructure, some

[30] The shortcomings of monetary and physical approaches to measuring infrastructure have been discussed in Chapter 3.

studies use data on the total length of highways, standardised by area to take account of differences in region size (e.g., Clark and Murphy, 1996; Duffy-Deno, 1998; Jiwattanakulpaisarn *et al.*, 2009a, 2009b; and Gebremariam *et al.*, 2011). On the empirical evidence, the literature using either monetary or physical measures of transport infrastructure obtains mixed results. Additionally, different types of transport infrastructure are found to have different impacts on employment.

Secondly, the modelling frameworks applied empirically to address the question of whether transport infrastructure affects employment can be classified into four categories. The most frequently-applied approach for estimating the employment impact of transport infrastructure is single-equation regression analysis. Empirical works based on this approach can then be divided into three main groups according to the type of data used for the analysis, namely, cross-sectional, time-series, and panel data. Several studies adopt a simultaneous-equation framework for employment and population growth (e.g., Boarnet, 1994; Clark and Murphy, 1996; and Duffy-Deno, 1998). Recently, Gebremariam *et al.* (2011) estimate a spatial panel simultaneous-equations growth model for employment, in-migration, out-migration, income, and government expenditure, showing a positive relationship between highways and employment growth. Instead of following the production function approach, some studies estimate the employment impact of transport infrastructure using a cost function (e.g., Seitz, 1993; and Nadiri and Mamuneas, 1998). A few studies apply a vector autoregression (VAR) technique to examine the empirical relationship between transport infrastructure and employment (e.g., Zografos and Stephanedes, 1992; and Jiwattanakulpaisarn *et al.*, 2009a).

Thirdly, variations with regard to the employment impact of transport infrastructure have emerged at both the aggregate level and sectoral level of empirical analysis. Much of the works in the empirical literature has been concerned with the impact of transport infrastructure on employment at the aggregate level. Some studies examine how transport infrastructure affects the overall employment, providing mixed results. Specifically, the relationship between transport infrastructure and employment (or, employment growth) is found to be positive and statistically significant in several studies (e.g., Jones, 1990; Crane *et al.*, 1991; Carroll and Wasylenko, 1994; Dalenberg *et al.*, 1998; Bollinger and Ihlanfeldt, 2003; and Gebremariam *et al.*, 2011), whereas others reveal no significant evidence or even a negative employment impact (e.g., Dalenberg and Partridge, 1995; and Jiwattanakulpaisarn *et al.*, 2009a, 2009b). Another group of empirical studies examines how transport infrastructure affects different sectors of the economy. Some of these studies focus on certain sectors, while others analyse the overall employment and then apply similar methodology to disaggregated employment data by industry. However, the empirical literature focusing on a specific industrial sector also provides mixed results. For instance, some studies find that manufacturing employment is positively associated with highway density, access to major highways, or public highway investment (e.g., Modifi and Stone, 1990; Crane *et al.*, 1991; Carroll and Wasylenko, 1994; and Bollinger and Ihlanfeldt, 2003); whereas, others provide no evidence of the impact of transport infrastructure on the manufacturing sector employment (e.g., Dalenberg and Partridge, 1995). Studies that analyse both total and sectoral employment suggest that the employment impact of transport infrastructure varies according to the sector concerned (e.g., Zografos and Stephanedes, 1992; and Carroll and Wasylenko, 1994).

Finally, few studies examine the spillover effect of transport infrastructure across geographical areas, providing evidence indicating that transport infrastructure installed in the neighbouring areas could produce either positive or negative spillovers to the employment in a particular area. Dalenberg *et al.* (1998) find a positive relationship between highways and employment growth at the U.S. state level; whereas, Haughwout (1999) argues that, according to the empirical evidence, infrastructure growth serves to redistribute employment growth from dense to less dense metropolitan-area U.S. counties. Also for the U.S. states, Jiwattanakulpaisarn *et al.* (2009a) show that the existence and the direction of spillovers vary according to the types of highway infrastructure and the time lags considered.

In sum, the employment impact of transport infrastructure has been studied widely, but the empirical question of whether transport infrastructure can generate job opportunities has remained inconclusive. Specifically, some studies find a positive employment impact of transport infrastructure while others show that the relationship between transport infrastructure and employment is not statistically significant or even turns out to be negative. A variety of modelling framework and econometric techniques has been applied, and the empirical analyses have been carried out at both aggregate and sectoral levels, though obtaining mixed results. With respect to the spillover effect of infrastructure, only a few studies consider whether infrastructure installed in the neighbouring areas could affect the employment in a particular area.

4.3 EMPIRICAL MODELS, VARIABLES AND DATA

4.3.1 Model Specification and Estimation Method

The reduced-form equation of private sector employment can be specified as a function of infrastructure and other factors potentially affecting labour demand and labour supply

$$\ln EMP_{i,t} = \alpha + \beta \ln ROAD_{i,t-1} + \gamma_{LD} LD_{i,t-1} + \delta_{LS} LS_{i,t-1} + \mu_i + \varphi_t + \varepsilon_{i,t} \quad (4.7)$$

where $EMP_{i,t}$ is the number of people working in the private sector in province i in year t; $ROAD_{i,t-1}$ is denotes the density of the road network, measured by total road length per 1,000 km^2, in province i in year $t-1$; $LD_{i,t-1}$ is the vector of other factors affecting the private sector labour demand in province i in year $t-1$; $LS_{i,t-1}$ is the vector of other factors affecting the labour supply in province i in year $t-1$; μ_i is the unobserved province-specific factors; φ_t is the unobserved time-specific factors; and $\varepsilon_{i,t}$ is the error terms.

Given the nature of the transition economy of Vietnam in which although economic reforms have led to the scaled-down restructuring of the SOE sector and the explosion of the private sector particularly in terms of the number of non-SOEs, SOEs still maintain a substantial role in the economy, the presence of SOEs in the labour market should be controlled for when estimating the determinants of private sector employment. Then, the employment model is re-written as

$$\ln EMP_{i,t} = \alpha + \beta \ln ROAD_{i,t-1} + \gamma_{LD} LD_{i,t-1} + \delta_{LS} LS_{i,t-1} + \sigma SOEEMP_{i,t-1} + \mu_i + \varphi_t + \varepsilon_{i,t} \quad (4.8)$$

where $SOEEMP_{i,t-1}$ is the proportion of people working in SOEs in total employment in province i in year $t-1$.

The present study examines further the employment impact of transport infrastructure by estimating the employment model which incorporates both own-province and neighbouring provinces' transport infrastructure

$$\ln EMP_{i,t} = \alpha + \beta \ln ROAD_{i,t-1} + \gamma_{LD} LD_{i,t-1} + \delta_{LS} LS_{i,t-1} + \sigma SOEEMP_{i,t-1} + \omega \ln w_{i,j} ROAD_{i,t-1} + \mu_i + \varphi_t + \varepsilon_{i,t} \quad (4.9)$$

where $w_{i,j}ROAD_{i,t-1}$ is the average road density of neighbouring provinces j of province i in year $t-1$.

One question arising from the estimation of a panel data model is whether the unobserved individual-specific effects, μ_i, and time effects, φ_t, should be treated as random variables or as fixed and estimable parameters. In the Fixed Effects (FE) model, the unobserved individual-specific effects and time effects can be controlled for, thus minimising the bias caused by omission of relevant variables. However, in the Random Effects (RE) model, it is assumed that μ_i, φ_t, and $\varepsilon_{i,t}$ are independent of each other and all other explanatory variables are independent of μ_i, φ_t, and $\varepsilon_{i,t}$ for all i and t.

In the present study, the null hypothesis of homoscedasticity is rejected in all specifications of the employment model. With respect to the problem of autocorrelation, the test statistic indicates that the errors follow first-order autoregressive process, AR(1), in all cases. The results of the tests for heteroskedasticity and autocorrelation are presented in Tables A1 and A2 in Appendix A.

The problem of AR(1) residuals can be solved by specifying a dynamic model with the inclusion of the lagged dependent variable as an explanatory variable. In that case, the most viable option should be the System-GMM estimator which can control for not only the endogeneity of variables, but also the problems of heteroskedasticity and autocorrelation. However, this chapter focuses on the long-run effects of infrastructure and other factors on employment and therefore does not consider the lagged dependent variable. In this case, the Feasible Generalised Least Squares (FGLS) estimator, which allows making specific assumptions about heteroskedasticity and/ or autocorrelation, is adopted to estimate the employment models specified above. Besides, all explanatory variables are lagged for one year to minimise the endogeneity problem, which may arise from such variables as infrastructure, and to support the assertion that employment is a predetermined decision.

4.3.2 Variables

Road density, i.e. road length per 1,000 km^2 of the provincial area, is used as a proxy for the availability of transport infrastructure. The present study assumes that an increased road density is an indicator of the road network improvement. Another assumption is that the quality of roads is homogenous at the province level. A positive sign is expected for the coefficient on the variable $ROAD$.

Several variables are used to capture the private sector labour demand which could be determined by the scale of production process, the accumulation of physical capital, or the growth of productivity. Those labour-demand variables include the volume of output, $OUTPUT$; the stock of physical capital, $CAPITAL$, or, alternatively, the physical capital stock per enterprise, $CAPITALPE$; the change of the Solow residual, $\Delta SOLOW$;[31] and the number of enterprises, $ENTERPRISES$. The coefficients on the output variable and the enterprise variable are expected to be positive. The sign of the coefficients on physical capital stock and productivity growth variables is left to be empirically determined. This is due to the uncertainty associated with the substitution and the complementary relationship between physical capital accumulation and labour demand, as well as between productivity growth and labour demand.

The empirical literature frequently uses demographic characteristics, such as the number of inhabitants in different ranges of ages, as variables controlling for the supply of labour. Due to the inconsistent availability of demographic data, the present study uses the number of higher-education institutions, $HIGEDU$, to measure the province's capacity to providing qualified labour, the change in the number of high school graduates, $\Delta SECGRA$, to measure the growth of the number of potential, new job seekers, the net migration rate, $NETMIG$, to control for an additional source of labour in-migrating from elsewhere, the urbanisation degree, $URBAN$, to measure the potential proportion of employees who are residing in the urban centres,[32] and the number of people residing in a particular province, $POPULATION$, to measure the potential size of labour pool. These proxies for the labour supply, especially the proxies for the supply of qualified labour, are expected to have positive impacts on private sector employment.

The inclusion of the proportion of SOE employment in total employment, $SOEEMP$, is to control for the SOE sector's presence in business activity, and particularly in the formal

[31] Following Dutta et al. (2006), the Solow residual is defined as $\ln A = \ln Y - \beta \ln K - \alpha \ln L$, where A is the Solow residual, Y is output, K is physical capital, L is labour, α and β the share of output going to labour and physical capita, respectively, α is defined as $\alpha = \frac{WL}{PY}$, W is the wage rate, P is the price index, and, $\beta = 1 - \alpha$.

[32] The present study assumes that, on average, urban residents would have more opportunities to access better education and other amenities and, therefore, their labour quality is higher than that of those living in the rural areas.

labour market, which might crowd out the private business activities and hence private sector employment. More precisely, the increased participation of SOEs in the formal labour market might intensify the competition for labour and other production inputs especially physical capital, thus affecting negatively the business activities and employment of the emerging private sector. The sign of the coefficient on this variable is thus expected to be negative.

Transport infrastructure is supposed to have an impact on employment not only in the area where it is installed but also in the neighbouring areas. The present study adds the average road density of neighbouring provinces into the employment model to examine the existence of infrastructure spillovers across provinces sharing administrative borders. The contiguity-based spatial weights matrix, w_{cont}, and the contiguity- and distance-based spatial weights matrix, w_{dist}, which have been described in Chapter 3, are used alternatively to calculate the road spillover variable, $wROAD$. The literature reviewed above suggests that the employment in a particular area can also be influenced either positively or negatively by infrastructure installed in its neighbouring areas. Then, the sign of the coefficient on the road spillover variable is left to be empirically determined.

4.3.3 Data

The panel dataset contains 61 provinces for 8 years. Due to the inclusion of the lagged difference variables, the time series dimension of the sample is reduced to 6 years and the total number of observations is 366. Tables 4.1 and 4.2 presents descriptive statistics and correlation matrix of the variables mentioned above. With respect to the sources of data, the road variable is calculated by using the data on road length provided by the Ministry of Transport and the data on area provided by the General Statistics Office of Vietnam; the data on the number of higher-education institutions and the number of high school graduates are provided by the Ministry of Education and Training; the remaining variables are calculated by using the relevant data provided by the General Statistics Office of Vietnam.

Table 4.1: Descriptive Statistics of Variables

Variable	Obs.	Mean	Std. Dev.	Min	Max
Log(EMP)	488	9.975	1.147	6.711	14.19
Log($WAGE$)	488	1.968	0.309	0.938	3.013
Log($ROAD$)	488	6.696	0.767	5.252	8.062
Log($w_{cont}ROAD$)	488	6.825	0.562	5.459	7.911
Log($w_{dist}ROAD$)	488	6.748	0.547	5.368	7.923
Log($OUTPUT$)	488	14.88	1.465	10.56	19.68
ΔLog($SOLOW$)	427	0.003	0.268	-2.248	1.999
Log($CAPITAL$)	488	13.70	1.538	8.824	18.99
Log($CAPITALPE$)	488	7.180	0.878	4.996	10.02
Log($ENTERPRISES$)	488	6.520	1.005	3.433	10.70
$NETMIG$	488	-0.740	5.473	-10.41	41.08
Log($URBAN$)	488	-1.656	0.552	-2.849	-0.158
Log($POPULATION$)	488	13.94	0.526	12.53	15.72
ΔLog($SECGRA$)	427	0.066	0.201	-1.450	1.347
$HIGEDU$	488	0.362	8.084	0	61
$SOEEMP$	488	0.399	0.190	0.024	0.907

Note: All monetary variables are deflated by provincial GDP deflator.

Table 4.2: Correlation Matrix of Explanatory Variables

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Log($ROAD$)	1.0000													
2	Log($w_{cont}ROAD$)	0.6805	1.0000												
3	Log($w_{dist}ROAD$)	0.7164	0.9742	1.0000											
4	Log($OUTPUT$)	0.4925	0.4471	0.4194	1.0000										
5	ΔLog($SOLOW$)	0.0883	0.0736	0.0539	0.1379	1.0000									
6	Log($CAPITAL$)	0.4673	0.4412	0.3948	0.9161	0.1216	1.0000								
7	Log($CAPITALPE$)	0.2808	0.3405	0.2742	0.5761	0.0708	0.7872	1.0000							
8	Log($ENTERPRISES$)	0.4696	0.3773	0.3644	0.8981	0.1242	0.8419	0.3299	1.0000						
9	$NETMIG$	0.1448	-0.0225	-0.0138	0.3719	0.0317	0.4051	0.2512	0.4003	1.0000					
10	Log($URBAN$)	0.0604	-0.1382	-0.1334	0.5210	0.0339	0.4788	0.2424	0.5208	0.4663	1.0000				
11	Log($POPULATION$)	0.3674	0.3915	0.3591	0.6807	0.1191	0.6397	0.2962	0.7199	0.1365	0.0973	1.0000			
12	ΔLog($SECGRA$)	-0.0617	-0.0621	-0.0568	-0.0499	-0.0246	-0.0489	-0.0299	-0.0487	0.0020	-0.0261	-0.0505	1.0000		
13	$HIGEDU$	0.3853	0.2064	0.2162	0.5329	0.0212	0.5477	0.1763	0.6840	0.5106	0.4688	0.5061	-0.0407	1.0000	
14	$SOEEMP$	-0.3017	-0.2934	-0.3059	-0.3832	0.0089	-0.3787	-0.2364	-0.3728	0.0048	0.0469	-0.0931	-0.0098	-0.0291	1.0000

Note: The sample contains 427 observations.

In addition to the issues relating to the infrastructure data discussed in Chapter 3, another data issue arises from the fact that the employment and wage variables do not have skill premium. Put differently, the employment data are not broken down into skilled and unskilled workers due to data limitations. Similarity, the wage data are the average real wages paid in the private sector, which do not contain disaggregate information on the wages payable to skilled and unskilled workers. This lack of skill premium of the data would affect the performance of some variables, particularly productivity-related variables, in the reduced-form employment equation discussed below and the reduced-form wage equation presented in Appendix B.

4.4 RESULTS

4.4.1 Results of the Employment Model

Results of the employment model are reported in Table 4.3. Wald χ^2 is reported as a diagnostic tool for the validity of the FGLS regression, which is statistically significant at the 1 per cent level in all cases. Also, the problems of heteroskedasticity and panel-specific first-order autocorrelation are controlled for in all regressions.

Table 4.3: FGLS Results of the Employment Model

	(1)	(2)	(3)	(4)	(5)
Log(*ROAD*) $_{i,t-1}$	0.193***	0.208***	0.204***	0.179***	0.152**
	(0.055)	(0.055)	(0.059)	(0.066)	(0.069)
Log(*OUTPUT*) $_{i,t-1}$	0.103***				
	(0.020)				
Log(*CAPITAL*) $_{i,t-1}$		0.013	0.014		
		(0.014)	(0.015)		
ΔLog(*SOLOW*) $_{i,t-1}$			-0.030**	-0.030**	-0.027*
			(0.014)	(0.014)	(0.014)
Log(*CAPITALPE*) $_{i,t-1}$				0.010	0.010
				(0.016)	(0.016)
Log(*ENTERPRISES*) $_{i,t-1}$				0.074***	0.088***
				(0.025)	(0.027)
NETMIG $_{i,t-1}$	0.002*	0.002	0.002*	0.002*	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log(*URBAN*) $_{i,t-1}$	0.354***	0.432***	0.439***	0.405***	0.381***
	(0.077)	(0.083)	(0.081)	(0.084)	(0.087)
ΔLog(*SECGRA*) $_{i,t-1}$	0.032	0.035*	0.040**	0.045**	0.046**
	(0.019)	(0.019)	(0.020)	(0.021)	(0.021)
HIGEDU $_{i,t-1}$	0.014***	0.012***	0.011***	0.009***	0.009***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Log(*POPULATION*) $_{i,t-1}$					-0.174
					(0.172)
SOEEMP $_{i,t-1}$	-1.047***	-1.125***	-1.054***	-0.977***	-0.930***
	(0.075)	(0.080)	(0.085)	(0.095)	(0.101)
Constant	9.310***	10.925***	10.982***	10.673***	13.321***
	(0.691)	(0.573)	(0.608)	(0.721)	(2.722)
Wald χ^2 [*p*-value]	602.18 [0.000]	449.08 [0.000]	369.71 [0.000]	385.66 [0.000]	388.70 [0.000]
Observations	366	366	366	366	366
No. of Provinces	61	61	61	61	61

Note:
1. The dependent variable is the number of people working in the private sector, Log(*EMP*).
2. All explanatory variables are lagged for one year.
3. Province and year dummies are included in all regressions.
4. Standard errors are in parentheses.
5. * p<0.100; ** p<0.050; *** p<0.010
6. The problems of heteroskedasticity and panel-specific first-order autocorrelation are controlled for in all regressions.

Performance of Transport Infrastructure Variable

Consider first the result for the employment impact of transport infrastructure which is proxied by the density of road network. The positive and statistically significant coefficient on the variable $ROAD$ indicates that provinces with higher road densities tend to have greater private sector employment than others. This result is robust and appears to not be sensitive to the other explanatory variables included in the employment model.

The magnitude of the coefficient ranges from 0.152 to 0.208, varying according to the model specification but being larger than that of labour demand variables and that of some labour supply variables. One explanation could be that the variable also captures the impact of transport infrastructure on the demand side and on the supply side of the labour market, as discussed previously in the literature review. The modest magnitude of the coefficient suggests that an increase of 1 per cent in the length of roads divided by 1,000 km^2 of the mainland area leads to an increase of 0.152 per cent in private sector employment.

Performance of Labour Demand Variables

The coefficient on the variable $OUTPUT$ is positive and statistically significant at the 1 per cent level, suggesting that an increase in private sector employment is driven by an expansion of the production process. The magnitude of the coefficient is 0.103, which indicates that a 1 per cent increase in private sector output causes an increase of private sector employment by 0.103 per cent.

In contrast, the statistically insignificant coefficients on the variables representing physical capital stock, $CAPITAL$, and physical capital stock per enterprise, $CAPITALPE$, indicates that physical capital accumulation does not necessarily lead to an increase in employment. The statistical weakness of the performance of physical capital variables can be attributed to the fact that these variables might capture competing effects. Specifically, an increased accumulation of physical capital stock could increase the number of labour required for the expanded production process with an application of new machines or technologies but, at the same time, it might cause some substitution away from labour given that the new technological application resulted in a transformation of the production process from labour-intensive to technology-intensive. However, it is worth noting that an increase in physical capital stock might not necessarily lead to an increase in productivity. This is because, for example, some enterprises would increase the number of cheap and simple machines used to make more products without considering the technological level of those machines, hence requiring more workers to operate the new machines but being unable to enhance the productivity of the production process.

The relationship between productivity growth, proxied by the change of the Solow residual, $\Delta SOLOW$, and employment is negative and statistically significant at the 5 per cent level or the 10 per cent level, varying according to the specification of the model. A growth of productivity has a negative impact on employment because a smaller amount of labour input is required to produce a given level of output demanded. The magnitude of the coefficient ranges from -0.027 to -0.030, which suggests that a 1 per cent increase in the measure of private sector productivity growth causes private sector employment to decrease by around 0.03 per cent.

Similarly to the performance of the output variable, the coefficient on the variable $ENTERPRISES$, which represents the number of enterprises in the private sector, is positive and statistically significant at the 1 per cent level. Assuming that the size of enterprises does not change, an increase in number of enterprises should lead to an increase in the number of employees. The magnitude of the coefficient is at around 0.08, suggesting that an increase of 1 per cent in the number of private sector enterprises leads to an increase of around 0.08 per cent in private sector employment.

As noted above, the composition of skilled and unskilled labour is not separated, which makes the interpretation of the relationship between productivity and employment more complicated. The lack of skill premium of the employment data could provide one explanation for the results showing that the physical capital variables, namely, $CAPITAL$ and $CAPITALPE$, do not act as important determinants of employment and why the employment effect of the productivity growth measured by $\Delta SOLOW$ is not as large as that of other labour demand factors, such as the volume of output and the number of enterprises. In particular, the performance of the variables $OUTPUT$ and $ENTERPRISES$ appears not to be affected by whether the dependent variable has skill premium or not.

Performance of Labour Supply Variables

As far as the labour supply factors are concerned, higher-education institutions, $HIGEDU$, which is used as a proxy for the capacity of a particular province in providing qualified labour, has a positive and statistically significant impact on private sector employment. One possible implication of this result is that the province with a higher level of tertiary education infrastructure endowment, and hence a higher capacity of qualified labour supply, tends to have a higher employment in the private sector. The magnitude of the coefficient ranges from 0.009 to 0.014, which suggests that a 1 per cent increase in the measure of the capacity in providing qualified labour causes an increase in private sector employment by at least 0.009 per cent.

The change in the number of high-school graduates, $\Delta SECGRA$, which is expected to represent for the change in regard to the supply of potential, new job-seekers, has a positive impact on private sector employment. However, the coefficient on this variable only turns out to be statistically significant at the 5 per cent level when the productivity growth variable enters into the regression. This implies that the explanatory power of the variable representing an increased supply of secondary-level qualified labour could also be influenced by whether private sector productivity is taken into account or not.

Turning to the results obtained for the potential source of labour supplied by immigrants, there is some evidence indicating a positive relationship between private sector employment and net migration rate, $NETMIG$. However, the impact of in-migration is not as strong as the impacts of other local attributes on private sector employment, as indicated by the magnitude as well as the statistical significance level of the coefficients on the respective variables. This might suggest that, on average, enterprises located in a particular province place a stronger importance on the local human capital resources rather than on the immigrants.

In comparison to other labour supply factors, the proportion of urban population has the largest employment effect. The coefficient on the variable $URBAN$ is positive and statistically significant at the 1 per cent level in all regressions. The magnitude of the coefficient ranges from 0.354 to 0.439, which indicates that an increase of 1 per cent in the proportion of people living in the urban centres leads to an increase of around 0.4 per cent in private sector employment. Assuming that the labour quality of urban residents is, on average, higher than that of those living in the rural areas, one interpretation of this result is that a higher proportion of productive labour is positively associated with an increase in private sector employment. In addition, the result for urbanisation might indicate that most of private sector enterprises should be concentrated in the urban centres, and therefore private sector employment turns out to be positively associated with the proportion of urban inhabitants in total population.

To some extent, urbanisation can have implications for, not only the supply side, but also the demand side of the labour market. An increase in the degree of urbanisation indicates an increase in the size of urban market relative to the overall size of the local market. Enterprises will earn more from the expansion of the urban market since the purchasing power of people living in urban centres is supposed to be higher than that of rural population. Then, an increased profitability may lead to an increase in the demand for labour so as to expand the production process.

The coefficient on the variable $POPULATION$ is not statistically significant, suggesting that the total number of people residing in a particular province does not affect private sector employment in that province. Indeed, the use of this variable as a proxy for the labour supply, or more precisely the potential size of labour pool, requires an assumption that the age structure is homogenous across provinces. Then, the statistical insignificance obtained for the variable could be attributed to the fact that the population variable does not provide information on the size of the labour force of each province.

Performance of Public Sector Variable

The variable proxying the ownership structure of the provincial economy is the proportion of SOE employment in total employment, $SOEEMP$, which is found to be negatively associated with private sector employment. The statistical significance of this employment impact falls below the 1 per cent level in all regressions. This negative relationship suggests that the deepening and accelerating of the scale-downward restructuring of the state sector, among other factors, has a significant explanatory power for the explosion of private entrepreneurship at least in terms of employment. The magnitude of the coefficient ranges from -0.930 to -1.125, which suggests that a 1 per cent increase of in the proportion of people working in SOEs causes the employment in the private sector to decrease by around 1 per cent.

$SOEEMP$ appears to be the main factor influencing private sector employment. This finding is in line with the salient facts presented in Chapter 2 regarding the on-going restructuring of the SOE sector and the explosion of the private sector over the sample period. Comparing all factors, infrastructure and urbanisation are the key determinants of an increase in private sector employment while the proportion of SOE employment in total employment in a particular province acts as the opposite force. Although a large SOE sector relative to the economy crowds out private sector employment as indicated by the empirical finding, some particular SOEs may also offer access to infrastructure services which are necessary for all activities, regardless of ownership, of the economy. On the other side, the crowding-out effect may be important, especially if the economic activities of SOEs are substitutable to those of the private sector. Then, a sectoral analysis would provide a further insight into the impact of the presence of SOEs in the economy on private sector employment, which is however unable to be performed in the present study due to the unavailability of data.

4.4.2 Results of the Employment Model with Road Spillover Variable

Results presented in Table 4.4 provide an illustration of the performance of the road spillover variable which is measured by using the contiguity-based binary spatial weights matrix and, alternatively, the contiguity- and distance-based spatial weights matrix. Similarly to the version without road spillover variable, the Wald χ^2 test is used as a diagnostic tool for the validity of the FGLS regression, which has the χ^2 statistic that is statistically significant at the 1 per cent level in all cases. Also, the problems of heteroskedasticity and panel-specific first-order autocorrelation are controlled for in all regressions.

Overall, there is some evidence of a negative cross-province spillover effect of the neighbouring road network. The negative coefficient indicates that the improvement in transport infrastructure in neighbouring provinces tends to draw job opportunities away from a particular province. Not only the magnitude but also the statistical significance of the coefficient on the road spillover variable, either $w_{cont}ROAD$ or $w_{dist}ROAD$, varies according to the set of other explanatory variables.

With respect to the performance of other variables as long as the road spillover variable enter into the regression, the road variable appears to experience the most significant change. Specifically, the inclusion of the road spillover variable reduces substantially the magnitude of the coefficient on the variable $ROAD$. This could be affected by the fact that the present study is unable to control for the connectivity between the road network of a particular province and the road network of its neighbouring provinces. However, the statistical importance of the road variable holds in all specifications of the employment model with the road spillover variable.

Table 4.4: FGLS Results of the Employment Model with Road Spillover Variable

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(*ROAD*) $_{i,t-1}$	0.147**	0.166***	0.183***	0.165**	0.148**	0.153***	0.159***	0.178***	0.158**	0.134*
	(0.060)	(0.060)	(0.061)	(0.067)	(0.071)	(0.057)	(0.058)	(0.061)	(0.066)	(0.071)
Log(w_{cont}*ROAD*) $_{i,t-1}$	-0.138	-0.230***	-0.201**	-0.233**	-0.214**					
	(0.084)	(0.089)	(0.094)	(0.093)	(0.100)					
Log(w_{dist}*ROAD*) $_{i,t-1}$						-0.099	-0.179**	-0.115	-0.165*	-0.138
						(0.072)	(0.081)	(0.087)	(0.091)	(0.098)
Log(*OUTPUT*) $_{i,t-1}$	0.091***					0.095***				
	(0.021)					(0.021)				
Log(*CAPITAL*) $_{i,t-1}$		0.024	0.019				0.022	0.018		
		(0.016)	(0.016)				(0.016)	(0.016)		
*Δ*Log(*SOLOW*) $_{i,t-1}$			-0.028*	-0.026*	-0.026*			-0.029**	-0.027*	-0.027*
			(0.015)	(0.015)	(0.015)			(0.014)	(0.015)	(0.015)
Log(*CAPITALPE*) $_{i,t-1}$				0.014	0.015				0.015	0.016
				(0.016)	(0.016)				(0.016)	(0.017)
Log(*ENTERPRISES*) $_{i,t-1}$				0.091***	0.098***				0.086***	0.095***
				(0.028)	(0.029)				(0.027)	(0.028)
NETMIG $_{i,t-1}$	0.003**	0.003**	0.003**	0.002**	0.002**	0.002**	0.002**	0.002**	0.002**	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log(*URBAN*) $_{i,t-1}$	0.379***	0.421***	0.426***	0.398***	0.387***	0.368***	0.422***	0.435***	0.396***	0.374***
	(0.085)	(0.093)	(0.089)	(0.091)	(0.095)	(0.083)	(0.092)	(0.088)	(0.090)	(0.095)
*Δ*Log(*SECGRA*) $_{i,t-1}$	0.038*	0.046**	0.046**	0.051**	0.050**	0.038*	0.046**	0.044**	0.049**	0.048**
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)
HIGEDU $_{i,t-1}$	0.014***	0.014***	0.013***	0.012***	0.011***	0.014***	0.014***	0.011***	0.011***	0.010***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Log(*POPULATION*) $_{i,t-1}$					-0.078					-0.122
					(0.191)					(0.190)
SOEEMP $_{i,t-1}$	-0.999***	-0.985***	-0.975***	-0.893***	-0.873***	-1.009***	-1.007***	-1.012***	-0.913***	-0.887***
	(0.083)	(0.089)	(0.090)	(0.100)	(0.103)	(0.082)	(0.090)	(0.090)	(0.101)	(0.104)
Constant	10.882***	12.562***	12.340***	12.138***	13.219***	10.471***	12.300***	11.891***	11.773***	13.477***
	(0.950)	(0.833)	(0.846)	(0.876)	(2.832)	(0.894)	(0.783)	(0.818)	(0.871)	(2.851)

Wald χ^2 [*p*-value]	399.22 [0.000]	315.17 [0.000]	296.11 [0.000]	317.66 [0.000]	318.02 [0.000]	490.30 [0.000]	344.43 [0.000]	315.00 [0.000]	321.71 [0.000]	320.82 [0.000]
Observations	366	366	366	366	366	366	366	366	366	366
No. of Provinces	61	61	61	61	61	61	61	61	61	61

Note: See Table 4.3

4.5 SUMMARY AND POLICY DISCUSSION

Main findings of Chapter 4 can be summarised as follows.

The variable of interest, i.e. the road variable, has a positive and statistically significant impact on private sector employment. This impact generally is not sensitive to what other explanatory variables are included in the regression. Regardless of the variance in the statistical significance of the road spillover variable across different specifications of the employment regression, there is some evidence indicating the existence of negative cross-province spillovers in the employment impact of roads. This could be driven by the possibility that transport infrastructure improvement alters the comparative advantages across provinces, thereby drawing economic activities and, in particular, jobs from elsewhere to the provinces where such the improvement takes place.

On the performance of the private sector's labour demand factors, there is evidence of a positive relationship between employment and the scale of the production process proxied by the output volume, as well as between employment and the number of enterprises located in the province, but a negative relationship between employment and productivity growth. Such the negative effect indeed is not very substantial in comparison to the effects of other factors on private sector employment. This could in part be explained by the lack of a skill premium in the dependent variable.

With respect to the impact of the labour supply factors on private sector employment, the capacity of qualified labour supply, proxied by the endowment of higher-education institutions, and the degree of urbanisation, proxied by the proportion of urban population in total population, have positive and statistically significant impacts while the statistical significance of the positive impacts of the growth of high-school graduates and the net migration rate vary according to the specification of the employment regression.

Last but not least, the proportion of SOE employment, which is found to be negatively associated with the employment in the private sector, appears to be the main factor influencing private sector employment. Caution, however, is required for any interpretation of the result. This is because the SOE and private sector employment variables do not contain information on the sectoral composition.

Policy implications of the main findings are discussed below.

Firstly, the effectiveness of transport infrastructure improvement as a policy instrument to stimulate private sector employment is supported by the empirical finding that an increase in the road density, i.e. the length of roads per 1,000 km^2 of the provincial area, leads to an increase in private sector employment. However, it is worth noting that the improvement of transport infrastructure examined in the present study is in physical terms, not monetary terms. Furthermore, the positive employment impact found for the road variable is subject to a number of assumptions which are essentially related to the quality of road infrastructure. Then, an improvement in the quality of infrastructure would make the economic returns of infrastructure improvement more substantial.

Nevertheless, employment gains from transport infrastructure in a particular province may come at the expense of other provinces as there is clear evidence of cross-province negative spillovers. In terms of policy implications, the infrastructure-related policies made at the provincial level should take into account the competitive pressure arising from the development of neighbouring transport networks. In addition to infrastructure-related policies, other targeted province-specific policies for the provinces with underdeveloped infrastructure systems are needed due to the existence of cross-province negative spillovers which might resulted in the fact that these provinces would be left further behind. From a central-levelled policy perspective, due to the network characteristics of transport inrastructure, the central government should pay special attention to the coordination of transport construction among provinces to minimise the adverse effect of the fact that, by altering investment patterns in transport infrastructure relative to those of the neighbouring provinces, each province has the ability to modify the size of its transport stock and hence achieve its targets, such as the employment target, at the expense of its neighbours.

Secondly, policies to facilitate the private sector either to expand its production process or to explode in the numeral terms should be an integral part of the set of policies aimed at stimulating employment. This is interpreted from the empirical results which suggest that either an expansion of the production process or an increased number of private sector establishments leads to an increase in employment.

Thirdly, the positive relationship between private sector employment and higher-education infrastructure endowment should be interpreted from the perspective of skilled-labour supply. The empirical result is not necessary to be supportive of an increase in the number of higher-education institutions in any particular province. Instead, an increase in the capacity of qualified labour supply would be beneficial for stimulating private sector employment. This argument is further supported by the finding on the positive relationship between employment and the growth of high-school graduates.

Fourthly, although the employment variable does not have a skill premium, the negative relationship between private sector employment and productivity growth could suggest that a proportion of unskilled workers might be left unemployed, and some of them might join in the informal sector of the labour market. Since productivity growth is important for the long-run development of the economy, policies that providing incentives are required for technological improvement. This in turn could lead to a replacement of unskilled labour by machines and skilled labour. Then, policies aimed at improving the quality of the labour force would contribute to controlling for the issue of unemployment facing unskilled and semi-skilled workers who have been a dominant force of the large labour pool of Vietnam.

Fifthly, it is important to take into account the phenomenon that the private sector economic activities should be concentrated in the urban areas since the degree of urbanisation is found to be positively related to private sector employment. As presented in Chapter 2, GSO (2011b) has noted the phenomenon of over-concentration in the urban areas. In this context, the congestion-related adverse effects arising from this concentration, especially in the longer-run, should be estimated when proposing policies aimed at supporting the expansion of economic activities of the private sector.

Finally, by showing that a decrease in the proportion of SOE employment leads to an increase in private sector employment, the present study provides an empirical illustration of the economic reforms which have been carried out over the sample period and resulted in the scaled-down restructuring of the SOE sector and an explosion of the private sector. Another relevant empirical illustration can be seen in the subsequent chapter which, although focusing on the role of infrastructure in shaping the location choices of private sector enterprises, shows that a decrease in the proportion of SOEs is positively related to an increase in new private sector establishments.

CHAPTER 5

INFRASTRUCTURE AND INDUSTRIAL LOCATION CHOICES

5.1 GENERAL BACKGROUND

From a firm-level perspective, as Porter (2000) argues, the location decision, through its impact on business costs and profits, can affect the development of a firm's competitive advantage. Also, the attributes of the area where the firm chooses to locate can have lasting effects on its performance. Therefore, from a policy perspective, understanding the local attributes that influence the location choices of firms is important not only for understanding the development potential of particular geographical areas, but also for making policies to promote further the development of private entrepreneurship, which in turn determines the development of the national economy as well as and the creation of job opportunities, and policies aimed at narrowing the development gap between the core and the periphery areas.

As Quinet and Vickerman (2004) summarise, the classical firm-location theories emphasise that firms locate in the area which enables them to minimise transport costs associated with both the supply of inputs and the delivery of outputs, but require a number of assumptions. Most notably, externalities are not controlled for within the classical firm-location framework, and the level and spatial distribution of demand are assumed to be given. Over time, as Preston (2001) notes, the emphasis has been less on transport costs in the classical sense, but increased on transport costs as a factor that interacts with agglomeration externalities. In addition, as suggested by McCann and Shefer (2004), for the understanding of the complex relationship between transport infrastructure and locational behaviour, it is necessary to examine conditions of agglomeration under which the reduction in transport costs due to transport infrastructure improvement will influence the movement of economic activities. Since the 1990s, the new economic geography literature has provided a further insight on industrial agglomeration which has different patterns across non-homogenous spaces. Specifically, the literature states that firms may concentrate in particular areas to take advantage of agglomeration economies. On the other side, the literature points out agglomeration diseconomies which essentially arise from competition and congestion. In this strand of literature, transport costs also have a role to play in explaining the locational behaviour of firms and transport infrastructure can accelerate the agglomeration process by improving accessibility to particular areas and connectivity between areas. Taking theoretical considerations from the classical firm-location theories and the new economic geography literature together could enhance the understanding of the geographical distribution of economic activities. It can be generalised that an improvement in transport infrastructure would shape the location choices of firms through its impact on reduced transport costs, which are one of the factors that affect the optimisation of the firm's needs and profits, and enhanced market access.

According to Hayter (1997), there are three approaches to identify determinants of locational behaviour, namely, the neoclassical, behavioural and institutional approaches. Factors determining the location choices of firms vary across these approaches. For the neoclassical approach, the location choices are determined by local attributes that affect profits such as transport costs, labour costs and agglomeration economies. In this sense, the neoclassical approach centres its location-choice analysis on profit-maximisation and cost-minimisation strategies. Unlike the neoclassical approach, the behavioural approach goes further by dealing with aspects left open by the earlier industrial theory, that is, imperfect information and uncertainty facing decision makers as well as firm heterogeneity in regard to sub-optimal location choices, industrial organisation and multi-goals. Finally, the institutional approach emphasises the need to understand the industrial location behaviour within a political economy framework. The majority of empirical studies have followed the neoclassical approach to explain the location choices of firms while those following the behavioural and structural approaches are rare (Arauzo-Carod *et al.*, 2010). Apart from other local attributes, transport infrastructure and market access, which can be assumed to be facilitated by transport infrastructure improvement, have been found to have a role to play in shaping the location choices of firms.

In Vietnam, one strikingly geographical disparity is the uneven distribution of economic activities across regions and provinces. Provinces with high concentration of enterprises are characterised not only by higher relative incomes, more industrialised, more open to external trade, but also higher endowment of infrastructure in comparison to less dynamic provinces. As presented in Chapter 2, the period 2000-2007 witnessed an explosion of the private sector, especially in terms of establishments. The state sector accounted for the average share of 6.62 per cent per year in the total number of enterprises, which was much smaller than the average share of 89.89 per cent per year of the non-state sector but about twice the average share of 3.49 per cent per year of the foreign-invested sector. While the non-state sector and the foreign-invested sector achieved high average rates of annual growth of 22.88 per cent and 18.49 per cent in their number of enterprises, respectively, the number of SOEs declined by an average rate of -6.83 per cent per year in the number of its enterprises. Given that, as shown in Chapter 2, the explosion of private sector establishments had been one dynamic aspect of the Vietnamese economy, but being distributed unevenly across geographical areas, during the period 2000-2007 and that, as found in Chapter 4, private sector establishments are positively associated with employment, which in turn is an important policy issue, over the same period, it is worth understanding why private sector enterprises choose to locate in a particular province, but not in others.

Chapter 5 does not try to find factors leading to the formation of new enterprises. Instead, using a panel dataset containing the information on the number of private sector enterprises, which first appear in the merged firm-level dataset, in the Vietnamese provinces over the period 2002-2007, the empirical analysis focuses on the impact of transport infrastructure, while also controlling for the impacts of other local attributes, on the location choices of these enterprises. The concept of local attributes, or province-specific characteristics, is extended beyond purely-natural features, such as geographical distances and natural resources. In particular, the locational effects of the following elements are examined: transport infrastructure proxied by the density of road network, which is measured as the total road length per 1,000 km^2 and is assumed to represent the availability of reliable infrastructure that reduces transport costs and enhances market access; local demand; wages and labour supply; agglomeration economies; and such public policies as the industrial zone policy and the SOE restructuring. Empirical results obtained in this chapter are expected to contribute to understanding how a province's non-purely-natural characteristics influence the location choices of new private sector enterprises, and to understanding further the role of infrastructure in explaining the geographical disparities in regard to the distribution of economic activities across provinces. Regarding methodology, this chapter adopts the count data modelling approach which, according to Guimarães *et al.* (2004), is helpful to understand why some particular provinces are chosen to be the location of enterprises while the others are not chosen by any enterprise. Due to the nature of the data used in the present study, the Negative Binomial Regression technique is applied as the main estimator for the count data regression model. In addition, by performing estimations using different sub-samples based on workforce size and industry, this chapter also shows that the location preferences of the private sector vary according to enterprise-specific characteristics. For the workforce-size based sub-samples, enterprises are divided into the four categorises of micro, small, medium and large enterprises. With respect to the sectoral sub-samples, this chapter focuses on manufacturing, trade and service sectors which, as presented in Chapter 2, had experienced a substantial increase in the number of establishments over the sample period.

The rest of Chapter 5 is organised as follows. Section 5.2 reviews the literature, focusing on the classical location-choice models, the agglomeration literature, and empirical studies on the locational effect of transport infrastructure. Section 5.3 presents the model specification and estimation methods, describes variables added in the empirical models, and discusses issues related to the data used in the empirical analysis. Results are presented and discussed in Section 5.4. Finally, Section 5.5 summarises this chapter and draws policy implications.

5.2 LITERATURE REVIEW

The literature review begins with a brief description of the classical location-choice models in which transport costs play a central role.[33] This is followed by a discussion on agglomeration economies, as well as the role of transport infrastructure in the process of industrial agglomeration, formulated in the new economic geography literature. Finally, empirical findings are summarised from previous econometric studies.

5.2.1 Classical Location-Choice Theories

The theory formulated by Von Thünen explains the optimal location for an agricultural producer, based on a strong assumption of perfect competition. The maximum rent that each agricultural producer has to pay for a given unit of land, $\hat{r}_i$, is determined by the following factors, as written in Vickerman (1980, p.49)

$$\hat{r}_i = q(p - c) - qtd_i \tag{5.1}$$

where q is output; p is the market price per unit; c is direct production costs; t is transport rate per unit of distance; d_i is distance from the market. The market is assumed to be "a single market" where there is perfect competition between producers; and, within the transport sector, there is no discrimination between users. Consequently, q, p, c and t are assumed to be constant. According to equation (5.1) and its assumptions, the maximum rent is determined uniquely by the location of the area where the producer is located. Specifically, transport costs, i.e. transport rate multiplied by output, qt, increase with distance from the

[33] According to Vickerman (1980), the main early developments in location theories can be found in the works of Von Thünen, Weber, Isard and Losch. However, Isard's analysis does not go much beyond the Weberian basic ideas (Vickerman, 1980). Therefore, the review of classical location-choice theories presented in this section will focus on the ideas of Von Thünen, Weber and Losch.

market, and the maximum rent payable by the producer decreases with increased transport costs. As Vickerman (1980) notes, equation (5.1) is a single-activity specification. Accordingly, the transport rate will change with an introduction of more than one activity into the model. Given that the transport rates for commodities are known, the relationship between r and d can be derived for each t, and the location choice of a particular producer is determined by the maximum rent, $\hat{r}_i$, payable at each d_i. Producers are divided into two groups which are expected to have contradictory locational behaviours: (i) those specialised in low-transport-cost activities and (ii) those specialised in high-transport-cost activities. The former would locate further away from the market while the latter would choose the areas closer to the market.

In the model of industrial location choice developed by Weber, the distribution of land is assumed to be uniform whereas the distribution of other production inputs, which are required in addition to land for most manufacturing industries, are not. Transport rates are assumed to be dependent on two factors: (i) the relative weights of inputs and outputs and (ii) the distance over which these inputs and outputs must be moved. The location choice of a manufacturing firm is determined by the relative pulls of various material locations and the market, and the question is whether the firm would locate nearer to the market or to the sources of materials. The basic Weber locational criterion is to minimise total transport costs, T, which are written in Vickerman (1980, p.51) as

$$T = tw_1 d_{f1} + tw_2 d_{f2} + \cdots + tw_n d_{fn} + td_m \quad (5.2)$$

where t denotes transport costs per ton-mile; w_1 … w_n are the weights of inputs $1 \ldots n$ required by per unit output; d_{f1} … d_{fn} are the distances that inputs 1 … n must be moved; d_m denotes the distance to the market. As can be seen in equation (5.2), d_{f1} … d_{fn} and d_m are not independent and, given the values of t and w_1 ... w_n, setting the value of any one distance will determine all the others. Since the price of the product and the prices of production inputs are assumed to be given and homogenous across areas, the optimal location involves finding a set of d_{f1} ... d_{fn} and d_m to minimise total transport costs, T. It is worth noting that the Weber theory focuses on multiple input sources and ignores outside demand and supply factors. According to the Weber theory, firms can be divided into two categories: (i) those who produce goods less heavy than the raw materials used in production and (ii) those producing heavier goods. The former would locate near to the raw-material sources while the latter would choose the areas near their output markets. Generally, the Weber model approaches the location problem of firms explicitly as a transport-cost minimisation problem.

The Losch model is claimed to be a considerable advance on the Weber model by moving to a world with demand and a world of areas rather than points (Vickerman, 1980). A number of important assumptions of the Losch model can be summarised as follows: (i) the existence of a homogeneous plain, (ii) the evenly and continuously distributed population, (iii) the population with identical preferences, (iv) the product's shipping cost paid by the consumer, (v) the existence of free entry into production, and (vi) the existence of rationality among consumers and producers. In regard to the final assumption, consumers would seek supply from the nearest producer while producers would attempt to maximise profits. These assumptions are referred to as "general conditions of equilibrium" (Parr, in McCann, 2002, p.34). With respect to the optimal location, the producer would choose the area where the net profit is greatest, and the net profit is defined as the difference between sales and production costs. Obviously, producers would prefer the area where that difference is greatest.

To sum up, in the classical theories, transport costs play a central role in shaping the location choices of firms. Accordingly, firms would prefer to locate in the area which enables them to minimise the costs of transporting inputs and outputs. As Quinet and Vickerman (2004) note, among flawed assumptions of the classical location-choice theories are the ignorance of externalities and the reliance on the given level and the given spatial distribution of demand. The New Economic Geography economists, such as Fujita *et al.* (1999) and Fujita and Thisse (2002), argue that the impractical view of self-balance in the relative growth of different areas has led to a considerably increased number of models exploring the organisation of a spatial economy. Essential features of those models are the increasing consumption tied to increasing wealth, the increasing returns to scale in production, and the cumulative process reinforced by agglomeration economies (Quinet and Vickerman, 2004). With respect to the role of transport infrastructure, McCann and Shefer (2004) suggest that since the costs of transaction facing firms have changed over time as a result of the evolution of technology, the conditions of agglomeration, under which the reduction of transport costs due to transport infrastructure improvement influences the movement of economic activities, should be examined so as to understand the complex relationship between transport infrastructure and locational behaviour.

5.2.2 Agglomeration and the Role of Transport Infrastructure

5.2.2.1 Economies of Agglomeration

From the input perspective, the classical location decision theories maintain that firms should minimise costs by locating near to the sources of production inputs. However, from the output perspective, classical theorists, in particular Hotelling (1929), argue that firms should minimise the proximity to their customers while maximising the distance from their competitors. Building upon this competition-based insight, the literature of industrial location has examined the phenomenon of agglomeration. According to McCann and Shefer

(2004), the evolution of the thought on agglomeration economies has centred on the notion of, and the benefits from, industrial clustering. Different types of industrial clustering, which take place at different industrial levels, will have different implications for associated agglomeration economies. Localisation economies and urbanisation economies are the two broad categories of agglomeration economies that are frequently mentioned.

Marshall (1920) is widely acknowledged to be the first to identify three sources of the economies of agglomeration: (i) knowledge spillovers among firms, (ii) geographically concentration of economic activities, and hence labour demand, that creates a pool of specialised labour, and (iii) concentration of economic activities, and hence production-input demand, that creates a pool of specialised input suppliers. Later, together with the contributions of Arrow (1962) and Romer (1986), the agglomeration literature often refers localisation economies, i.e. cost-saving externalities that are maximised when a local industry is specialised, to the Marshall-Arrow-Romer externalities. Given that an industry is subject to Marshall-Arrow-Romer externalities, firms are likely to locate in a few areas where producers specialised in that industry have already concentrated (Lall *et al.*, 2003).

If localisation economies arise within a particular industry, urbanisation economies occur across all industries. According to Lall *et al.* (2003), the Chinitiz-Jacobs diversity, proposed by Chinitiz (1961) and Jacobs (1969), implies that the industrial diversity is important for knowledge spillovers, which primarily occurs across industries, and their associated beneficial externalities. More precisely, the diversity of knowledge sources concentrated and shared in urban centres, or cities, make these areas become breeding grounds for new ideas and inventions, and hence new products. Lall *et al.* (2003) also note several benefits of economic diversity that go beyond the knowledge-spillover argument. Specifically, firms locating in larger and more diverse cities will have larger home markets for skilled workers and end products as well as a better access to such business services as banking, financial and legal advisers, and advertising and real estate services. More importantly, the externalities of economic diversity can stem from the heterogeneity of economic activities, or the variety of goods available for consumption and goods being produced, in the large and diverse cities (Lall *et al.*, 2003). Put differently, in those cities, the consumption utility can be enhanced with an increased range of local consumer goods while the production level can be raised with an increased range of local producer goods.

Opposing the elements that favour the geographical concentration of economic activities, the literature also identifies a number of centrifugal forces. For instance, the benefits of agglomeration may be offset by the costs associated with the increased competition for such production inputs as labour and land, which in turn causes wages and rents to rise, and the higher transport costs due to congestion. These agglomeration diseconomies may deter firms from locating in the areas where economic activities have already concentrated. However, Burgess and Venables (2004) argue that, particularly in developing countries, agglomeration economies could outweigh diseconomies and, therefore, make it beneficial for firms to locate in large cities.

5.2.2.2 Determinants of Agglomeration

The question in regard to the spatial distribution of economic activities has been examined in several strands of literature, such as residential and firm location theories, urban economics and, especially, the new economic geography. Two frequently-cited main approaches to explain the geographical concentration of economic activities include the Sachs approach, or the "first nature" approach, and the Krugman approach, or the "second nature" approach. According to Chasco and López (2010), the "first nature" factors include natural features exogenous to the economy, such as climate or natural resources; whereas, agglomeration economies represent the "second nature" factors.[34] Chasco and López (2010) also note that the question as to how much geography remains a matter of importance for economic development seems to be inconclusive and, additionally, it is likely to be difficult to identify the net effect of the "first nature" factors on the "second nature" factors since there might be "a close connection" between the two categories.

With respect to the connection between the "first nature" and the "second nature", on the one side, as Chasco and López (2010) note, the "first nature" geography is viewed as an initial advantage that can be amplified by the "second nature" agglomeration economies. On the other side, as pointed out by Krugman (1993), the "first nature" advantages tend to create the "second nature" advantages through the cumulative processes. However, the Krugman (1993) also argues that agglomeration could be explained by the "second nature" agglomeration economies alone. In the same vein, Fujita *et al.* (1999) argue that, for some productive activities, the only key factor explaining the spatial inequality of economic activity distribution could be the increasing returns to scale.

According to McCann and Shefer (2004), the new economic geography literature has been developed since the seminal paper of Krugman (1991), which generally explains the spatially uneven distribution of economic activities as a result of market processes under the conditions of agglomeration economies. In particular, Krugman (1991) analyses the performance of two main agglomeration forces, including (i) economies of scale, which are considered as the attractive force, and (ii) transport costs, which are considered as the centrifugal force, within a two-region modelling framework to explain why and when the

[34] As noted by Chasco and López (2010), the "first nature" factors are also called "pure geography" (Henderson, 1999) and the "second nature" factors are also called "man-made" economies.

concentration of manufacturing firms occurs in a few areas, leaving other areas relatively undeveloped.

The model of Krugman (1991) is based on a number of assumptions. First of all, the economy is divided in two regions: region 1 and region 2; and, there are two types of products in this economy: agricultural products, A, and manufacturing products, M. The agriculture sector is the constant-returns sector while the manufacturing sector is the increasing-returns sector. With respect to the market structure of that economy, there is monopolistic competition among suppliers of manufacturing intermediates while competitive behaviour exists in the agriculture sector. Secondly, in regard to consumption, all individuals share the same utility function and the aggregate consumption of manufactured products is defined as the constant elasticity of substitution (CES) composite of manufacturing intermediates. Thirdly, labour is the only endowment factor, and there are two categories of workers. The first category contains workers that are mobile across regions, accounting for a fraction φ of the world population;[35] whereas, the second category contains agricultural workers who are immobile between region 1 and region 2, with a given supply of $(1-\varphi)/2$ in each region. Fourthly, labour is the only input used to produce agricultural products; whereas, the production of a manufacturing product involves a fixed cost and a constant marginal cost.[36] The fixed cost is considered as a source of economies of scale. Finally, while transportation of agricultural products is costless, the costs of transporting manufacturing products take the Samuelson's "iceberg" form.[37] Due to the costless transportation, the price of agricultural products and,

[35] The composite of the fraction μ is defined as $L_1 + L_2 = \varphi$, where L_1 and L_2 denote the manufacturing labour supply in region 1 and the manufacturing labour supply region 2, respectively.

[36] The production of a manufacturing product is specified as $x_i = \frac{L_{Mi}}{\beta} - \frac{\alpha}{\beta}$, where L_{Mi} is the number of manufacturing workers used in producing the manufacturing product i, and x_i is the manufacturing product's output, and $L_{Mi} > \alpha$. Under the assumption of free entry, in each region, all manufacturers produce the same regardless of the wage rate, relative demand and so on.

[37] Specifically, only a fraction $\tau < 1$ of each unit of manufacturing products transported from region 1 to region 2 (or, from region 2 to region1), arrives in region 2 (or, region 1).

then, the earnings of each agricultural worker in region 1 are the same as those of each agricultural worker in region 2.

As Krugman (1991) notes, whether regions converge or diverge is determined by the three key parameters, namely, the share of income spent on manufacturing products, μ, the elasticity of demand, σ, and the inverse index of transport costs, τ. The main ideas of Krugman (1991) can be summarised as follows. In the two-region model, a country can be differentiated into a core and a periphery. While the core is assumed to be industrialised with the concentration of manufacturing firms, the periphery is the home of agricultural activity. The emergence of a core-periphery pattern depends on transport costs, economies of scale, and the importance of the manufacturing sector in the national economy. Particularly with respect to the influence of transport costs, Krugman (1991) shows that firms tend to locate in regions with larger market demand to realise the economies of scale and minimise the costs of transportation. For the purpose of stating the ideas of Krugman (1991) in relation to the role of transport costs, the below mathematical description is drawn directly from his paper.

In the short-run equilibrium, total income of manufacturing workers of each region is equal to total spending, specified for region 1 as

$$w_1 L_1 = \mu \left[\left(\frac{z_{11}}{1 + z_{11}} \right) Y_1 + \left(\frac{z_{12}}{1 + z_{12}} \right) Y_2 \right] \tag{5.3}$$

and specified for region 2 as

$$w_2 L_2 = \mu \left[\left(\frac{1}{1 + z_{11}} \right) Y_1 + \left(\frac{1}{1 + z_{12}} \right) Y_2 \right] \tag{5.4}$$

and the distribution of manufacturing workers and their wages will determine the incomes of the two regions, specified for region 1 as

$$Y_1 = \frac{1-\mu}{2} + w_1 L_1 \tag{5.5}$$

and specified for region 2 as

$$Y_2 = \frac{1-\mu}{2} + w_2 L_2 \tag{5.6}$$

where z_{11} is the ratio of region 1's spending on local manufacturing products to its spending on imported manufacturing products; z_{12} is the ratio of region 2's spending on local manufacturing products to its spending on imported manufacturing products; w_1 and w_2 are wages in region 1 and region 2, respectively; L_1 and L_2 denote the manufacturing labour supply in region 1 and the manufacturing labour supply region 2, respectively; and Y_1 and Y_2 are the incomes of region 1 and region 2, respectively.

Since transport costs are assumed to incur in the manufacturing products, they affect the demand and then the total spending. This can be seen in the equation specified for z_{11}

$$z_{11} = \left(\frac{n_1}{n_2}\right)\left(\frac{p_1\tau}{p_2}\right)\left(\frac{c_{11}}{c_{12}}\right) = \left(\frac{L_1}{L_2}\right)\left(\frac{w_1\tau}{w_2}\right)^{-(\sigma-1)} \tag{5.7}$$

and the equation specified for z_{12}

$$z_{12} = \left(\frac{n_1}{n_2}\right)\left(\frac{p_1}{p_2\tau}\right)\left(\frac{c_{11}}{c_{12}}\right) = \left(\frac{L_1}{L_2}\right)\left(\frac{w_1}{w_2\tau}\right)^{-(\sigma-1)} \tag{5.8}$$

and the relative demand equation for representative manufacturing products

$$\frac{c_{11}}{c_{12}} = \left(\frac{p_1\tau}{p_2}\right)^{-\sigma} = \left(\frac{w_1\tau}{w_2}\right)^{-\sigma} \tag{5.9}$$

where c_{11} is the consumption in region 1 of a representative manufacturing product produced in region 1, and c_{12} is the consumption in region 1 of a representative manufacturing product produced in region 2; p_1 is the free-on-board price of the local manufacturing product, and $p_2\tau$ is the price of the imported manufacturing product; n_1 and n_2 are the number of manufacturing products produced in region 1 and region 2, respectively.

In the long-run equilibrium, transport costs are present in the price index of manufacturing products, specified for the inhabitants of region 1 as

$$P_1 = \left[f{w_1}^{-\sigma(\sigma-1)} + (1-f)\left(\frac{w_2}{\tau}\right)^{-(\sigma-1)}\right]^{-1/(\sigma-1)} \tag{5.10}$$

and specified for the inhabitants of region 2 as

$$P_2 = \left[f\left(\frac{w_1}{\tau}\right)^{-\sigma(\sigma-1)} + (1-f){w_2}^{-(\sigma-1)}\right]^{-1/(\sigma-1)} \tag{5.11}$$

Complete agglomeration is another case considered in Krugman (1991). Assuming that all manufacturing workers and, hence, all manufacturing production are concentrated in region 1, the market size of region 1 will therefore be larger than that of region 2. Since a share of total income is spent on manufacturing products, μ, and all this incomes goes to region 1, the region 1's manufacturing production, $Y_1 - Y_2 = w_1 L$, has to serve the demand in both regions, $\mu(Y_1 + Y_2)$. Then, the value of sales of the region 1 firm, V_1, is specified as

$$V_1 = \frac{\mu}{n}(Y_1 + Y_2) \tag{5.12}$$

where n is the total number of manufacturing firms.

In the context that all manufacturing workers are concentrated in region 1, a firm must pay higher wages to attract workers so as to produce in region 2. This firm is referred to as the "defecting" firm. The wages paid in region 2 relative to those in region 1 are defined as

$$\frac{w_2}{w_1} = \left(\frac{1}{\tau}\right)^{\mu} \tag{5.13}$$

Due to the higher wages, the marginal cost of producing in region 2 and, hence, the price of manufacturing product produced in region 2, are higher than those in region 1. Then, the value of sales of the "defecting" firm vis-à-vis the region 1 firm is rescaled by $(w_2/w_1\tau)^{1-\sigma}$ when selling to region 1 and $(w_2\tau/w_1)^{1-\sigma}$ to region 2. Hence, the value of sales of the "defecting" firm, V_2, is specified as

$$V_2 = \frac{\mu}{n}\left[\left(\frac{w_2}{w_1\tau}\right)^{1-\sigma} Y_1 + \left(\frac{w_2\tau}{w_1}\right)^{1-\sigma} Y_2\right] \tag{5.14}$$

In equation (5.14), transport costs, τ, capture the "defecting" firm's advantage in selling in region 2, but its disadvantage in selling to region 1. The ratio of the value of sales of the "defecting" firm to the value of sales of the region 1 firm, which can be derived from equations (5.12), (5.13) and (5.14), v, has the form

$$v = \frac{V_2}{V_1} = \frac{1}{2}\tau^{\mu(\sigma-1)}[(1+\mu)\tau^{\sigma-1} + (1-\mu)\tau^{1-\sigma}] \tag{5.15}$$

The above specification of v does not control for the fixed cost. Because the condition of the zero profit implies $V_i \propto w_i\alpha/\beta \equiv$ fixed cost (where i represents either region 1 or region 2), a profitable deviation has to satisfy $V_2/V_1 > w_2/w_1 = \tau^{-\mu}$. This reduces to the analysis of $v > 1$, and then v is re-written as

$$v = \frac{V_2}{V_1} = \frac{1}{2}\tau^{\mu\sigma}[(1+\mu)\tau^{\sigma-1} + (1-\mu)\tau^{1-\sigma}] \tag{5.16}$$

From equation (5.16), Krugman (1991) identifies a set of parameter values shaping the boundary between concentration and non-concentration.[38] First of all, both a larger share of income spent on manufacturing product (i.e. larger μ, which also implies a relatively larger size of the region 1 market) and a higher elasticity of substitution (i.e. higher σ, which also implies smaller economies of scale) suggest that it is unprofitable for the "defecting" firm to produce in region 2 when all other manufacturing production are concentrated in region 1.

[38] $\partial v/\partial \mu$, $\partial v/\partial \tau$ and $\partial v/\partial \sigma$ are calculated as

$$\frac{\partial v}{\partial \mu} = v\sigma(\ln\tau) + \frac{1}{2}\tau^{\sigma\mu}[\tau^{\sigma-1} - \tau^{1-\sigma}]$$

$$\frac{\partial v}{\partial \tau} = \frac{\mu\sigma v}{\tau} + \frac{\tau^{\mu\sigma}(\sigma-1)[(1+\mu)\tau^{\sigma-1} - (1-\mu)\tau^{1-\sigma}]}{2\tau}$$

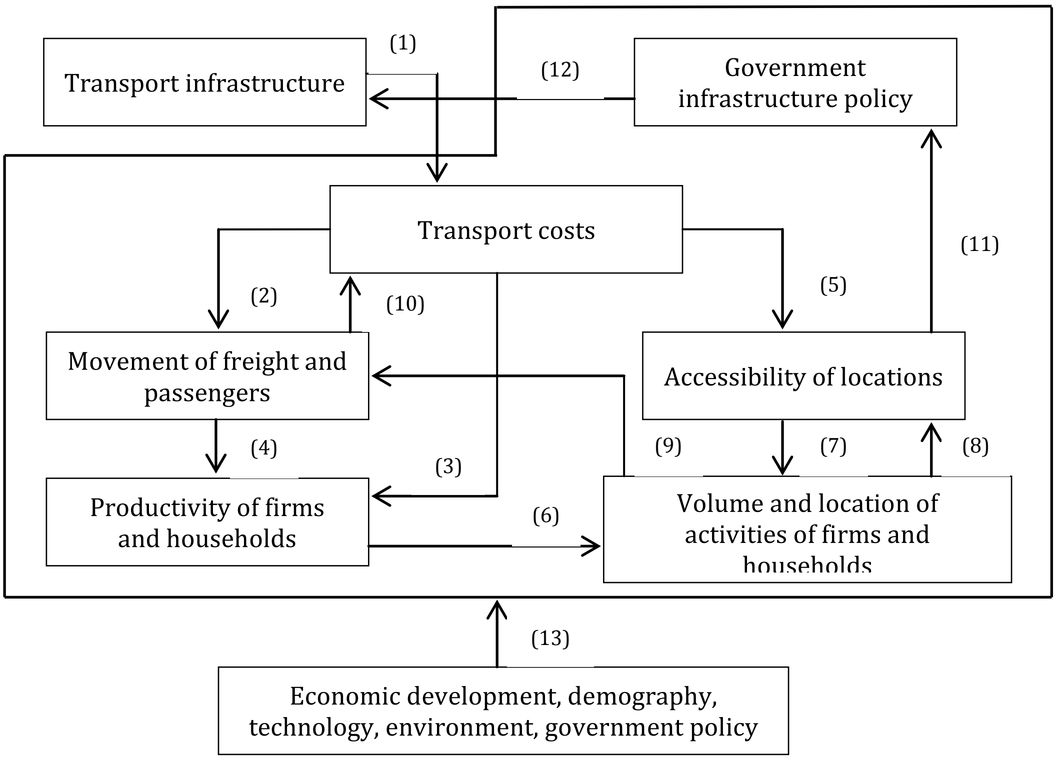
$$\frac{\partial v}{\partial \sigma} = \ln(\tau)\left\{\mu v + \frac{1}{2}\tau^{\mu\sigma}[(1+\mu)\tau^{\sigma-1} - (1-\mu)\tau^{1-\sigma}]\right\} = \ln(\tau)\left(\frac{\tau}{\sigma}\right)\left(\frac{\partial v}{\partial \tau}\right)$$

With respect to transport costs, there are three cases. When transport costs are equal to zero (i.e. $\tau = 1$), location is irrelevant (i.e. $v = 1$). When transport costs are so high (i.e. low τ) that the ratio v exceeds one (i.e. $v > 1$), it is profitable for the "defecting" firm to set up its manufacturing production in region 2 despite the fact that all other manufacturers concentrate in region 1. In contrast, when economies of scale are so large (i.e. small σ) or the share of income spent on manufacturing product is so large (i.e. large μ) that the ratio v falls below one (i.e. $v < 1$), it is unprofitable for the "defecting" firm to set up its manufacturing production in region 2, regardless of how high the costs of transportation are, as long as all other manufacturers concentrate in region 1.

5.2.2.3 The Role of Transport Infrastructure

Bruinsma (1995) proposes a conceptual framework explaining how transport infrastructure affects the location of economic activities.

Figure 5.1: Transport Infrastructure and the Spatial Pattern of Economic Activities



Source: Bruinsma (1995)

The central element of the model is that transport infrastructure improvement leads to a reduction in transport costs, which in turn produce three major effects: (i) an increase in the accessibility of particular areas, (ii) an increase in the productivity of firms and households, and (iii) a change in the volume of production and the concentration of firms and households. Aside from these three major effects, of particular importance is the fact that the impact of transport infrastructure on the spatial pattern of economic activities is also influenced by economic development, demographic, technological and environmental factors, and public policies, as can be seen in Figure 5.1. This in turn implies that the endowment of transport infrastructure is necessary, but not sufficient, to shape the location choices. From the spatial perspective, if transport infrastructure does lead to an increase in the concentration of economic activities in some particular areas, there may be other areas where this concentration is reduced; and, this can be explained by redistribution forces associated with the improvement of accessibility. Although at the national level, due to an increase in the overall level of transport infrastructure provision, and thereby a reduction in aggregate transport costs, the positive effects can be expected to outweigh the negative for the economy as a whole, the wider disparities between, for example, the core and the periphery areas within a country can be intensified by an improvement in transport infrastructure.

As Bruinsma and Rietveld (1996) define, accessibility is “the potential of opportunities for interaction” between economic agents in different parts within a particular area and between economic agents in that particular area and those locating in other areas. From an inter-regional perspective, Vickerman (1996) argues that the relationship between transport infrastructure and accessibility centres on the level of connectivity, which refers to a chance for firms locating in a particular area to develop profitable relations with firms, and/or consumers, locating in the other areas. Accordingly, a high level of connectivity would enhance the linkages between firms across geographical areas. On the other side, an inadequate endowment of transport infrastructure would result in a lack of connectivity which, in turn, implies a lack of opportunities for establishing inter-regional profitable relations.

According to Fujita and Thisse (2002), the new economic geography literature hypothesises that transport infrastructure improvement could not only encourage concentration but also cause dispersion of economic activities. In essence, this would depend on the level of initial transport costs, the amount of reduced transport costs, and the extent of agglomeration economies (e.g., Venables and Gasiorek, 1999; Vickerman, 2002). Alañón-Pardo and Arauzo-Carod (2011) describe the “twofold” relationship between improved accessibility and agglomeration economies as follows. On the one hand, the geographical scope of agglomeration economies may be enlarged by accessibility improvement. This would arise from an increase of opportunities for interaction when economic agents are closer in terms of their geographical distance. On the other hand, an improved accessibility may cause production to be moved to the periphery and, then, goods produced in the periphery are transported to the core's market. This geographical dispersion of economic activity would lead to an erosion of the benefits of agglomeration, as also argued by Haughwout (1999).

From a perspective of firm heterogeneity, the evaluation of the relationship between locational behaviour and transport infrastructure should also take into account issues relating to firm specific characteristics which, in turn, could affect their transportation demand. The survey of empirical studies provided in the subsequent section will clarify this point further. As McCann and Shefer (2004) demonstrate, there is an association between firm heterogeneity and each type of industrial clusters, namely, the pure agglomeration, the industrial complex, and the social network. What is important is the existence of these different clustering types has different implications for the understanding of the role of transport infrastructure in shaping the location choices of firms (McCann and Shefer, 2004).

5.2.3 Econometric Evidence on the Locational Effect of Transport Infrastructure

There is an extensive empirical literature seeking to identify factors explaining why firms choose a particular area to locate their factories. Hayter (1997) suggests a broad categories of those location factors: (i) transport facilities, (ii) materials, (iii) markets, (iv) labour, (v) external economies, i.e. urbanisation and localisation, (vi) energy, (vii) community infrastructure, i.e. social overhead capital and economic overhead capital, (viii) fixed and financial capital, (ix) land/ buildings, (x) environment, i.e. amenity and policy, and (xi) government policy. Indeed, there is no one way of classifying local attributes, or location-specific characteristics, which vary according to the geographical areas concerned and produce either direct effects or indirect effects on the locational behaviour of firms. Furthermore, the set of factors determining the industrial location choices can be different for different firms, for firms specialised in different industries, or for firms at different stages of development.

With respect to transport infrastructure, the empirical literature assumes transport infrastructure improvement as a factor that attracts the location choices of firms by increasing the productivity and hence the profitability of doing business in a particular area. The empirical studies surveyed in this sub-section can be divided into two groups: those explicitly estimating variables proxying for infrastructure, such as road density, and those estimating variables measuring geographical distance to transport infrastructure. The latter appears to outnumber the former. In addition, all of the studies surveyed in this section examine the locational effects of agglomeration economies, among other local attributes.

Following the count data modelling approach, Holl (2004) estimates the locational effect of transport infrastructure, among other factors, on firm birth in the industry and service sectors in Portuguese municipalities over the period 1986-1997. In this study, the market potential index is used to control for differences in the costs of transporting between particular municipalities and the main market centres, and the motorway access dummies are expected to capture the advantages of locating near to the new motorway corridors. The results of the Negative Binomial Regression model indicate that the new transport infrastructure has a statistically significant impact on firm birth. As the author explains, the new instalment of transport infrastructure enhances the attractiveness of nearby municipalities to new establishments. The study also shows that there are differences among sectors in regard to the extent to which the new transport infrastructure affects firm birth.

Concerning the results for other municipality-specific variables, firm birth is encouraged by an increased local demand, a more diversified industrial structure of the local economy and a better-qualified workforce. The evidence of the locational effect of agglomeration economies stemming from sectoral specialisation is relatively weak. With respect to the impact of wages, an increase in the wage level affect negatively firm birth in the manufacturing sector but have little impact on the service sector. Overall, from the study of Holl, we can see that there is evidence indicating a positive locational effect of new transport infrastructure though the effect varies according to the sectors observed.

Also for Portugal, Holl (2005) examines determinants of the spatial distribution of manufacturing plant start-ups and plant relocations over the period 1986-1997 using both the Poisson Regression model and the Negative Binomial Regression model. Transport infrastructure is proxied by the motorway access indicator which is constructed by calculating the distance between each municipality and the nearest inter-regional motorway. The results show that the locational effects of local attributes vary according to whether the firm is a new establishment or it moves from elsewhere. Specifically, proximity to inter-regional motorways matters for both new establishments and re-locations, but the locational effect of proximity is greater for the latter category of firms in comparison to the former one. With respect to other local attributes, new establishments show a greater interest in the municipalities having larger local market size, higher labour force qualification, lower labour costs and a more diverse economic environment. On the contrary, relocations are more attracted by better national market accessibility, more producer services and a larger industrial basis. Generally, the study of Holl shows that the extent to which transport infrastructure, as well as other local attributes, determines the location choices of firms depends on the characteristics firms, i.e. whether they are start-ups or relocating ones.

Also for Portuguese manufacturing firms but for the period 1992-2007, Mota and Brandão (2011) adopt the count data modelling approach to examine the differences in the sets of determinants of the locational behaviour of single-plant and multi-plant firms. The study hypothesises that the location choices of new establishments would be determined by supply factors (land, labour and capital costs, workforce and technological characteristics), demand factors (market size and market accessibility) and agglomeration economies. The results show that factors such as urbanisation economies, land costs and the size of the local market would influence the location choices of new multi-plants. For single plants, they are more sensitive to the costs of labour, the economies of localisation and urbanisation, and the distance to main markets while making their location choices. The study, however, does not explicitly investigate the role of transport infrastructure. Instead, the locational effect of a variable measuring the geographical distance from each municipality to important cities of Portugal is estimated, and the results indicate that a greater distance exerts a negative effect on the probability that the municipality is chosen by firms to locate. One important conclusion from the study of Mota and Brandão is the differences concerning the location choice preferences of multi-plants versus single plants to agglomeration economies.

For Chinese provinces, Li and Park (2006) examine factors determining the location choices of foreign firms across different industries in the years 1999 and 2002. The dependent variable measures the ratio of the number of new foreign firms in a given industry entering a given province during the period of 1999 and 2002, to the number of domestic firms in a given industry entering a given province before 1999. Put differently, the dependent variable is the ratio of new foreign firms during the period 1999-2002 to existing domestic firms before 1999 in the same industry in a province. The location choices of foreign firms are explained by three categories of location factors, namely, infrastructure, agglomeration economies, and institutional changes. Infrastructure variables include consumption of electricity per capita, telephone lines per capita, and road length per capita. The results show that foreign firms tend to choose to locate in provinces with better infrastructure. Concerning the results for other variables, localisation of foreign firms and domestic firms has the strongest impact on foreign location choices. Specifically, new foreign firms tend to go to the provinces with a high concentration of existing foreign firms but are less likely to choose the provinces with a high concentration of domestic firms. The result for the size of the local market, proxied by provincial GDP, is statistically insignificant; whereas, there is evidence of a positive effect of institutional changes. Overall, one noteworthy aspect of this study is that it differentiates between domestic and foreign agglomeration in regards to their locational effects on foreign firms.

Cheng (2006) is one example of the study adopting the Conditional Logit model and conducting the Hausman-McFadden test on whether the Independence of Irrelevant Alternatives (IIA) assumption holds.[39] The study, however, focuses on the locational effect of the labour cost rather than that of transport infrastructure. In this study, transport infrastructure, proxied by the total length of highways and railways divided by the area, has no statistical importance in determining the location choices of Japanese firms in Chinese provinces.

Chang *et al.* (2010) explore the location choices of Japanese and Taiwanese multinational enterprises (MNEs) in Chinese provinces between 1996 and 2005, using the Conditional Logit model and Nested Logit model. The development of infrastructure is measured by the length of highway. The results provide no evidence indicating the locational effect of highway infrastructure. Indeed, this study focuses on the role of agglomeration of same-nationality

[39] The estimation of the Conditional Logit model relies heavily on the IIA assumption. Whether the IIA assumption holds in the Conditional Logit model can be examined by the Hausman-McFadden test. The Conditional Logit model is discussed further in the subsequent section.

firms in the locational behaviours of MNEs. For the locational effect of agglomeration, it is generally concluded that (i) MNEs tend to invest in Chinese provinces where agglomeration of same-nationality firms exists, (ii) less productive Japanese firms prefer to locate close to larger same-nationality agglomerations, and (iii) there are no differences in location according to firms' productivity in the case of Taiwanese firms. Although this study does not focus on explaining the role of transport infrastructure as a location factor, it explores an interesting aspect regarding agglomeration economies, that is, the agglomeration of same-nationality firms. However, this study does not report whether IIA assumption holds in the Conditional Logit model.

Another example of the locational effect of transport infrastructure can be seen in the study of Tokunaga and Jin (2011) which adopts the Negative Binomial Regression model for count data to examine the role of highway infrastructure in explaining the location choices of Japanese manufacturing firms in Chinese provinces. The results show that the higher the density of highway installed in a particular province, the greater the number of Japanese manufacturers choosing to locate in that province. In addition, the study employs three measures of market potential: domestic market potential, the number of ports, and distance from each Chinese province to Tokyo, Japan, and finds that the larger domestic market potential and the increased availability of ports will increase the probability that a province is chosen by Japanese manufacturers while a greater geographical distance will lead to a decrease of that probability. Also, in this study, there is evidence of a positive locational effect of same-nationality-agglomeration.

On determinants of the location choices of Spanish manufacturing firms, Alañón-Pardo and Arauzo-Carod (2011) adopt a Negative Binomial Regression model in which the set of explanatory variables includes road accessibility, municipality-specific characteristics as well as characteristics of neighbouring municipalities. The results show that agglomeration economies and accessibility play an important role in shaping the location choices of firms. As argued by the authors, accessibility improvement increases the attractiveness of a certain municipality as well as the potential interaction between municipalities. Such the attractiveness can also be explained by localisation economies, urbanisation economies or the "follow the leader's location decision" strategy. Concerning the spatial spillover effects, the authors assume that industrial activity outside a certain municipality has an effect on the activity within that municipality. This point is proved by some econometric evidence on the spillover effects of human capital and agglomeration across neighbouring municipalities, although the significance of these effects varies according to the specialised industry of the firm. Overall, the study of Alañón-Pardo and Arauzo-Carod (2011) emphasises the importance of taking into account the spatial implications for studying the locational behaviour of firms, and it provides evidence indicating that the spillovers generate different effects according to different industries where the firms specialise.

After all, the most important conclusion arising from the above survey of empirical evidence is that the impacts of local attributes on the location choices of firms vary according to firm-specific characteristics, such as firm size, industrial activity, or industrial organisation. To examine the potentially important implications of transport infrastructure improvement for the distribution of firms across geographical areas, previous studies implicitly assume that an improvement in transport network leads to a reduction in transport costs and an enhancement of market access. Some studies find evidence supporting the role of transport infrastructure in explaining the location choices of firms, while others provide results indicating that transport infrastructure is less likely to be as important as other local attributes, such as labour costs and agglomeration economies. Recently, Alañón-Pardo and Arauzo-Carod (2011) find evidence suggesting the importance of considering the spatial implications of several local attributes for explaining the spatial distribution of economic activities. While analysing the locational behaviour, it is necessary to recall the conclusion of Blair and Premus (1987) on the different implications of different location factors in different stages of economic development. That is, because the economy needs to shift to advanced technologies to remain competitive, the importance of non-traditional location factors, such as business climate or labour skills, would increase while the traditional location factors, such as transport costs, labour and access to raw materials, may decline in their influences on the decisions of firms to locate in a particular area. Similarly, Hayter (1997) concludes that the location choices of firms vary according to their specific preferences to production inputs, infrastructure, and other geographically heterogeneous and immobile factors, but greater emphasis has been given to the importance of public policies, business strategies and structures, and the structure of labour markets, as well as to the interactions between these factors.

5.3 EMPIRICAL MODELS, VARIABLES AND DATA

5.3.1 Model Specification and Estimation Method

Econometric research on the location choices of firms usually adopts the discrete choice model that relies on the Random Utility Maximisation (RUM) framework proposed by McFadden (1974). In the discrete choice framework, the location-choice probability is modelled in a partial equilibrium setting where a representative firm maximises its profits subject to uncertainty deriving from unobservable location-specific characteristics. In this model, the dependent variable is a binary variable taking the value of one if the representative firm chooses to locate in a particular province i and the value of zero if not. An alternative is the count data modelling approach. In the count regression model, the dependent variable is the number of firms choosing to locate in each province. Both approaches have their own pros and cons for analysing the industrial location choices.

5.3.1.1 The Conditional Logit and Nested Logit Models

Within the discrete choice framework, a profit function is used to explain why the representative firm chooses to locate in a particular area. Mathematically, the profit π of the representative firm n in province i can be written as

$$\pi_{n,i} = X_{n,i}\beta' + \varepsilon_{n,i} \tag{5.17}$$

where X refers to the vector of the observable characteristics of the province; β' is the vector of coefficients on the variables X; and ε is the independently and identically distributed error terms.

Province i is chosen by the representative firm n if and only if

$$\pi_{n,i} > \pi_{n,I}, \forall i \neq I \tag{5.18}$$

where I indexes all possible location-choice alternatives, i.e. provinces, available to the representative firm n.

The probability of the representative firm n choosing a particular province i out of I potential provinces can be expressed as

$$\text{Prob}(n, i) = \frac{\exp(X_{n,i}\beta')}{\sum_{s=1}^{I} \exp(X_{n,s}\beta')} \tag{5.19}$$

where $\text{Prob}(n, i)$ is the probability that the representative firm n chooses to locate in province i; and the estimates of β' can be obtained by using the maximum likelihood method, as demonstrated in Greene (2008a).

Let $d_{n,i} = 1$ if the representative firm n chooses to locate in province i, and $d_{n,i} = 0$ otherwise, the log likelihood of the Conditional Logit model can be written as

$$\log L_{cl} = \sum_{n=1}^{N} \sum_{i=0}^{I} d_{n,i} \log \text{Prob}(n, i) \tag{5.20}$$

The Conditional Logit model is widely applied for analysing the discrete choice because the choice-probability form of the model is convenient and the globally concave likelihood function of the model allows the maximum-likelihood estimation to be straightforward (Heiss, 2002). However, in the Conditional Logit model, the error terms, which capture the unobserved characteristics that might determine the choices, are assumed to be independently and identically distributed. In other words, the Conditional Logit estimation is based on the assumption that the relative choice-probabilities between alternatives, i.e. provinces, depend only on the attributes of the alternative observed, not on those of other available alternatives. This is referred to as the Independence of Irrelevant Alternatives (IIA) assumption. If there are unobserved province-specific characteristics that affect the choice of both, for example, province A and province B similarly, the error terms of these provinces are correlated. Consequently, the IIA assumption is violated and, therefore, the Conditional Logit estimation provides biased results.

It is worth noting that, as McFadden (1974) suggests, the IIA assumption implies that the Conditional Logit model should be applied to analyse the discrete choice only where the choice-alternatives "can plausibly be assumed to be distinct and weighed independently in the eyes of each decision maker." In practice, the IIA assumption is often violated in the location choice analysis (e.g., Cheng, 2006; and Nefussi and Schwellnus, 2010). This is because, as explained by Nefussi and Schwellnus (2010), the location-choice alternatives that have similar income levels or locate within broader geographical regions tend to be more similar among themselves than other alternatives. This leads to the move to the models that relax or do not require the IIA assumption.

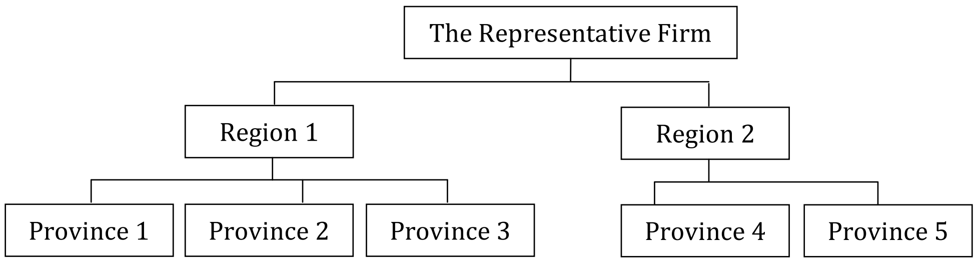
The Nested Logit model is viewed as an alternative to the Conditional Logit model, which is based on a partition of the set of choice-alternatives into sub-groups (Greene, 2008a).[40] These

[40] Heiss (2002) notes that there are several types of the Nested Logit models and a slight difference in the specification of the outcome probabilities of those models can lead to substantial differences in the results. One of these models is the Random Utility Maximisation Nested Logit model, which is viewed as preferable in most

sub-groups are also referred to as nests. In the Nested Logit model, the location choice of the representative firm n is assumed to be a hierarchical process. Let provinces be categorised into different groups of region, the probability of choosing a particular region, r, can be expressed as a function of the attributes of that r and the attributes of all provinces belonging to that r; whereas, the probability of choosing a particular province, i, can be expressed as the result of the probability of choosing a particular region, r, where i is nested and the probability of choosing i is conditional on having chosen that r. However, no testing procedure has been proposed for the specification of the nest structure, and this is viewed as a problematic aspect of the Nested Logit model (Greene, 2008a). In the same vein, Nefussi and Schwellnus (2010) note that one drawback of the Nested Logit model is associated with the arbitrariness in the choice of the nest structure.

Assuming that a country has 5 provinces which are divided into 2 regions, Figure 5.2 provides an illustration of the nest structure of location choice.

Figure 5.2: Hierarchical Location Choice Structure



Assuming a two-level choice structure, in which all of the alternatives, i.e. provinces, I, are divided into R nests, i.e. regions, $r_1, r_2 \dots r_k \dots r_m$, the description of the Nested Logit probability is given below

situations. An alternative is the Non-Normalised Nested Logit model. See Heiss (2002) for a detailed discussion on the differences between these models.

$$\text{Prob}(n,i) = \text{Prob}(n,r_k)\text{Prob}(n,i|n,r_k) \tag{5.21}$$

where

$$\text{Prob}(n,r_k) = \frac{\exp\left(Z_{n,k} + \lambda_{n,k} IV_{n,k}\right)}{\sum_{m=1}^{R} \exp\left(Z_{n,m} + \lambda_{n,m} IV_{n,m}\right)} \tag{5.22}$$

and

$$\text{Prob}(n,i|r_k) = \frac{\exp(X_{n,i}/\lambda_k)}{\sum_{j\in r_k} \exp\left(X_{n,j}/\lambda_k\right)} \tag{5.23}$$

where IV is the inclusive value of the nest r, and λ is the inclusive value coefficient of the nest structure, specified as

$$IV_{n,k} = \ln \sum_{j\in r_k} \exp(X_{n,j}/\lambda_k) \tag{5.24}$$

The parameters in $\text{Prob}(n,i)$ are estimated by using the maximum likelihood method, as can be seen from the demonstration presented in Greene (2008a). According to McFadden (1978), the inclusive value coefficient represents the degree of independence between alternatives, i.e. provinces, within the same nest, i.e. region, and its value needs to be in the range of between zero and one to be consistent with RUM. The statistically significant coefficients of the inclusive values would indicate that the hierarchical structure of location choice is highly relevant (Hansen, 1987).

5.3.1.2 The Poisson and Negative Binomial Regression Models

One of the most frequently-used techniques for modelling count data is Poisson Regression (Greene, 2008b). According to Guimarães *et al.* (2003), the Poisson Regression model finds its roots in the RUM framework but takes advantage of the equivalence between the Conditional Logit and the Poisson Regression models, which allows us to overcome a potential IIA violation. The Poisson probability function for province i receiving a number of firms N is specified as

$$\text{Prob}(N_i|X_i) = \frac{\exp(-\lambda_i)\lambda_i^{N_i}}{(1 + N_i)} \tag{5.25}$$

where N_i is the number of firms choosing to locate in province i, β' is the vector of estimated parameters, X_i is the vector of province-specific characteristics, and $\lambda_i = \exp(\alpha + \beta' X_i)$. As noted by Greene (2008b), the log-linear conditional mean function, $E(N_i|X_i) = \lambda_i$, and the equidispersion, $\text{Var}(N_i|X_i) = \lambda_i$, are the two features of the Poisson Regression model.

One important assumption of the Poisson Regression model is that the conditional mean λ_i is equal to the conditional variance. Put differently, in the Poisson Regression model, the variance of the random variable is constrained to equal the mean. However, in practice, the variance is frequently larger than the one assumed by the Poisson Regression model. This is referred to as over-dispersion, which causes downward biased estimates of the standard errors. In the location choice analysis, the over-dispersion in the dependent variable can be explained by the concentration of firms in a few geographical areas. The Negative Binomial Regression model, which arises as a gamma mixture of Poisson distributions, can be adopted as an alternative to the Poisson Regression model to overcome the problem of over-dispersion (Greene, 2008b).

According to Greene (2008b), the Negative Binomial Regression model relaxes the equidispersion restriction of the Poisson Regression model, which is motivated through the introduction of latent heterogeneity in the conditional mean of the Poisson Regression model. This is illustrated by, as demonstrated by Greene (2008b),

$$E(N_i|X_i) = \exp(\alpha + \beta' X_i + \varepsilon_i) = h_i \lambda_i \tag{5.26}$$

and

$$f(h_i) = \frac{\theta^\theta \exp(-\theta h_i) h_i^{\theta-1}}{\Gamma(\theta)}, h_i \geq 0, \theta > 0 \tag{5.27}$$

where $h_i = \exp(\varepsilon_i)$ is assumed to be a one parameter gamma distribution, $G(\theta, \theta)$ with mean 1 and variance $1/\theta = \kappa$, and ε_i is the error terms.

Then, the Negative Binomial distribution is obtained by integrating h_i out of the joint distribution

$$\text{Prob}(N_i|X_i) = \frac{\Gamma(\theta + N_i) r_i^\theta (1 - r_i)^{N_i}}{\Gamma(1 + N_i)\Gamma(\theta)} \tag{5.28}$$

where $r_i = \theta/(\theta + \lambda_i)$; and, the parameters of this model can also be estimated by the maximum likelihood method demonstrated in Greene (2008a).

As Guimarães *et al.* (2004) argue, although the count data modelling approach is useful to know not only how many times a province was chosen by firms, but also which provinces were not chosen by any firm, one unfavourable aspect of the count regression model is that the model lacks a sound microeconomic foundation of profit maximisation compared with the discrete choice model.

Hausman *et al.* (1984) propose Fixed-Effects estimations of the Poisson Regression and the Negative Binomial Regression which are aimed at accounting for unobserved time-invariant heterogeneities that can affect the location choices of firms. Furthermore, in the context of an analysis of the locational effect of transport infrastructure, Holl (2004) suggests the adoption of fixed-effects estimations which can control for the casual links between firm location choice and infrastructure construction. More precisely, as argued by Holl (2004), since the construction of new infrastructure is designed to connect important economic areas, it is necessary to take into account the fact that these areas have had higher numbers of firms and better systems of infrastructure simultaneously. However, it is important to note that, in the study of Holl (2004), transport infrastructure is proxied by dummies of distance from new motorways, which is then regressed on the number of new establishments; whereas, in the present study, transport infrastructure is represented by a stock variable which is also lagged for one year so as to minimise the problem of endogeneity. In addition, it is worth recalling that, as noted previously, in Vietnam, both economic centres and lagging economic regions could be targets for an increased provision of infrastructure. Furthermore, over the sample period, the majority of new roads were not constructed in the economic centres which, however, witnessed an explosion in the number of new establishments.

On the one side, Allison and Waterman (2002) argue that the conditional Fixed-Effects Negative Binomial Regression model proposed by Hausman *et al.* (1984) is not a “true fixed effects” model for count data since it does not control for all time-invariant covariates. Instead, they suggest an unconditional Negative Binomial Regression model that uses dummy variables to control for unobserved individual effects. In the same vein, Guimarães (2008) shows that, in the analysis of count panel data, the individual fixed effects might not be removed by the conditional maximum likelihood estimation of the Negative Binomial Regression model with fixed effects. On the other side, Cameron and Trivedi (1998) note the incidental parameters problem arising from using the dummy variable approach for the panel with a short time dimension, which could make the estimates of the parameters inconsistent.

In general, the Poisson Regression and Negative Binomial Regression models differ from each other in regards to their assumptions of the conditional mean and the conditional variance of the dependent variable. The choice between these two models should be based on the nature of the distribution of the dependent variable. Specifically, whether the distribution of the dependent variable is over-dispersed and hence should be estimated by the Negative Binomial Regression technique instead of the Poisson Regression technique can be examined through the Pearson χ^2 goodness-of-fit test, which can be incorporated along with the Poisson Regression model. In particular, if the statistical significance of the Pearson χ^2 statistic falls below the standard threshold of 0.05, the Negative Binomial Regression model is more appropriate for the dataset employed.

5.3.1.3 Model Specification and Estimation Method

Chapter 5 adopts the count data modelling approach to analyse the location choices of newly-established private sector enterprises. The specification and the results of the discrete choice model are reported in Appendix C, which generally suggest that the count regression model is referable. The choice between the Poisson Regression and the Negative Binomial Regression models is decided by considering the nature of the distribution of the count variable of new enterprises. As presented in Appendix A, the result of the Pearson test suggests that the Negative Binomial Regression model is more appropriate for the present study.

The base specification of the count regression model is written as

$$N_{i,t} = \alpha + \beta \ln ROAD_{i,t-1} + \delta_D D_{i,t-1} + \mu_L L_{i,t-1} + \psi_A A_{i,t-1} + \omega_G G_{i,t-1} + \gamma_r \sum_{r=1}^{5} R_r + \omega_t \sum_{t=1}^{5} Y_t + \varepsilon_{i,t} \quad (5.29)$$

where $N_{i,t}$ is the count of new enterprises belonging to the private sector in province i in year t; $ROAD_{i,t-1}$ denotes the density of the road network, measured by total road length per 1,000 km^2, in province i in year $t-1$; $D_{i,t-1}$ is the vector of variables capturing local demand in province i in year $t-1$; $L_{i,t-1}$ is the vector of variables capturing wages and labour supply in province i in year $t-1$; $A_{i,t-1}$ is the vector of variables capturing agglomeration economies in province i in year $t-1$; $G_{i,t-1}$ is the vector of variables capturing the implementation of public policies in province i in year $t-1$; R_r is the vector of region dummies; Y_t is the vector of year dummies; and $\varepsilon_{i,t}$ is the error terms.

Instead of applying the Conditional Fixed Effects Negative Binomial Regression estimation method, the present study uses the Negative Binomial Regression model in which region dummies are introduced to account for unobserved time-invariant heterogeneities. Year dummies are also introduced in order to control for time-related unobserved characteristics. All explanatory variables are lagged by one year to allow a response time by enterprises as well as to mitigate the possible problem of endogeneity. Note that it is the region dummies, not province dummies, introduced into the model. This is because the purpose of the inclusion of these region dummies is to account for unobserved time-invariant location heterogeneities, including natural advantages and disadvantages, or the "first nature" factors discussed in the literature review. The standard classification of Vietnamese provinces into six regions, as noted in Chapter 1, can capture those natural characteristics. In addition, the use of region dummies is aimed at avoiding the problem that there can be an overlapping between variables capturing province-specific characteristics which are time-invariant for some provinces, for example, the higher-education infrastructure variable, and province dummies, and especially between the industrial zone variable, which is both zero-valued and time-invariant for the majority of provinces, and province dummies.

The empirical analysis presented in this chapter also examines further the locational effect of transport infrastructure by estimating the following model which incorporates both own-province and neighbouring provinces' transport infrastructure

$$N_{i,t} = \alpha + \beta \ln ROAD_{i,t-1} + \delta_D D_{i,t-1} + \mu_L L_{i,t-1} + \psi_A A_{i,t-1} + \omega_G G_{i,t-1} + \Omega \ln w_{i,j} ROAD_{i,t-1} + \gamma_r \sum_{r=1}^{5} R_r + \omega_t \sum_{t=1}^{5} Y_t + \varepsilon_{i,t} \quad (5.30)$$

where $w_{i,j}ROAD_{i,t-1}$ measures the average road density in the neighbouring provinces of province i in year $t-1$.

Workforce-size based and sector based sub-samples are used to examine whether the locational effects of infrastructure and other local attributes hold in different samples of enterprises. This is also aimed at testing the hypothesis that the location preferences of enterprises could vary according to their specific characteristics. Basically, enterprises are divided into four categories of workforce size: micro, small, medium and large. For sectoral analysis, the present study focuses on five sectors: machinery manufacturing, non-machinery manufacturing, wholesale, retail, and services (including finance, insurance, property and other business services). To facilitate the sectoral analysis using sector-based sub-samples, not only the dependent variable, but several explanatory variables are also calculated at the province-sector level. These sub-samples are described further in the section on data issues.

5.3.2 Variables

Road density, i.e. road length per 1,000 km^2 of the provincial area, is used as a proxy for the availability of transport infrastructure. The present study assumes that an increased road density is an indicator of the road network improvement. Another assumption is that the quality of roads is homogenous at the province level. Given that the evaluation of the relationship between infrastructure and location choices is much more complex than simply a matter of an increased endowment of infrastructure, transport infrastructure improvement is assumed to be positively associated with an increased access to the local market. A positive sign is expected for the coefficient on the variable $ROAD$.

Population is used to measure local market demand. Enterprises located in provinces with larger markets are assumed to have more opportunities to supply their products to a larger number of consumers at lower transport costs and hence gain higher profits. Additionally, an increase in population leads to not only an increased demand for goods and services but also an enlarged pool of labour; both of which should attract the location choices of enterprises. Therefore, the coefficient on the variable $POPULATION$ is expected to have a positive sign. Another variable proxying the local demand is the growth rate of GDP, $GDPGR$, which captures the overall economic condition of the provincial economy and is also expected to have a positive locational effect.

The vector of variables proxying for the local labour market includes the real average wage level, $WAGE$, which is measured as the total wage bill divided by the total number of employees; the capacity of qualified labour supply, which is measured by the endowment of higher-education infrastructure, $HIGEDU$; and the net migration rate, which is viewed as a proxy for an additional source of workers and customers, $NETMIG$. The empirical relationship between the wage level and the locational behaviour could be ambiguous. If wages were considered as a measure of production-input costs, a location would become less attractive because of an increase in its wage level. On the contrary, higher wages would not dissuade enterprises that benefit from externalities associated with being located in large cities where wages would be higher than elsewhere. In addition, given that a higher wage level could reflect an increased labour quality, the higher costs caused by higher wages would be defrayed by higher productivity. However, due to the inconclusiveness of the relevant empirical literature, particularly at the sector level, the present study leaves the locational effect of the variable $WAGE$ being empirically determined. With respect to the locational effect of the capacity of qualified labour supply, the present study hypothesises that enterprises looking for a better educated workforce could prefer to locate in provinces with a higher endowment of higher-education institutions. It is also necessary to assume that provinces with a higher endowment of skilled labour would be preferred, even if this implies higher wages. The coefficient on the variable $HIGEDU$, i.e. the number of higher-education institutions, is then expected to have a positive sign. Regarding the variable $NETMIG$, the sign of the coefficient is also expected to be positive. This is based on the assumption that the entry of migrants into a particular province would raise the labour supply as well as the market demand within that province, which would in turn attract the location choices of enterprises. However, whether the coefficient is statistically significant would be dependent on whether the enterprises place an importance on the in-migration as an additional source of workers and customers. Besides, the performance of this variable should be affected by the fact that in-migrants would contribute to intensifying the congestion problem and thus congestion-related diseconomies, particularly in large cities.

The literature also claims that localisation and urbanisation economies influence the locational behaviour, as reviewed previously. Localisation economies are external to the enterprise but internal to the industry, while urbanisation economies are external to both the enterprise and the industry. Following previous studies, the formulas used to calculate indices of localisation and urbanisation economies are specified as

$$LOCALISATION_{i,s,t} = \frac{E_{i,s,t}}{E_{i,t}} \Bigg/ \frac{E_{61i,s,t}}{E_{61i,t}} \tag{5.31}$$

$$HERFINDAHL_{i,t} = \sum_{j=1}^{J} \left(\frac{E_{i,j,t}}{E_{i,t}} \right)^2 \tag{5.32}$$

where E is the number of employees; i, s, j and t represent province, economic sector, 2-digit industry and year, respectively; $61i$ and J denote total number of provinces and total number of industries, respectively. In the present study, for the purpose of empirical analysis, the 2-digit industries are divided into the following categories: (i) machinery manufacturing, (ii) non-machinery manufacturing, (iii) wholesale, (iv) retail, (v) services (including finance, insurance, property and other business services), and (vi) the sector containing the rest of 2-digit industries. The sector (vi) will not be examined in the empirical analysis presented below.

The localisation indices are calculated using complete information of all provinces and of all enterprises belonging to all categories of ownership, i.e. the state sector, the non-state sector and the foreign-invested sector. While the measure of specialisation focuses on a single industry, the Herfindahl index considers the industry mix of the entire provincial economy. The largest value for the Herfindahl index is 1 when a single industry dominates the economy. A higher Herfindahl index indicates a less industrial diversity of the economy; whereas, a lower index indicates a greater industrial diversity. In the present study, the calculation of the Herfindahl index uses complete information of all enterprises, regardless of their ownership, belonging to all 2-digit industries of a particular province. As Holl (2004) notes, the literature has emphasised the importance of agglomeration economies in shaping the industrial location choices, but there is an on-going debate on whether specialisation or diversity provides a more attractive business environment. The economies and diseconomies arising from agglomeration leave us with an ambiguous expectation about the sign of the coefficients on the variables $HERFINDAHL$ and $LOCALISATION$.

The average size of enterprises, $ESIZE$, is measured by the average number of employees per enterprise. The coefficient on this variable is expected to be positive under the assumption that the development of entrepreneurship, measured by the average workforce size, also implies a dynamic and competitive business environment where these enterprises are operating. The variable might also capture the overall development of local amenities facilitating business activities.

The vector of public-policy variables includes the industrial policy, IZ, and the SOE-restructuring policy, $SOECAP$. The on-going industrialisation is a geographically uneven process, and the industrialisation process in Vietnam should have been accelerated by, among other factors, the establishment of industrial zones (IZs). The variable capturing the effect of this industrial policy is constructed as

$$IZ_{i,t} = \frac{IZAREA_{i,t}}{ENT_{i,t-1}} \tag{5.33}$$

where $IZ_{i,t}$ represents the implementation of industrial policy, i.e. the policy to promoting industrialisation through industrial zone development, in province i in year t; $IZAREA_{i,t}$ is the total area, in hectare, of all industrial zones (including export processing zones) in province i in year t; and $ENT_{i,t-1}$ is the total number of enterprises located in province i in year $t-1$. When this variable enters into the regression as an explanatory variable, it will also be lagged for one year in the same way applied to other variables.[41] The rationale for the specification of the variable IZ is given below.

Firstly, the total number of industrial zones is not used as a proxy for industrial policy because it cannot capture the province-level variations in regard to the scope of industrial zones. Secondly, by divided the total land area of all industrial zones located in a particular province, which is assumed to represent the IZs supply, by total number of enterprises located in that province, which is assumed to represent the IZs demand, this specification also aims at controlling the possibility of an over-supply of industrial zones which in turn might imply an inefficient use of resources. Here, it is assumed that it would take about one year to complete the construction of an industrial zone and, therefore, the one-year lagged demand is assumed to be the demand at the time when the zone was designed. In practice, it would take more than one year to complete the construction. However, due to the short-time dimension of the panel dataset, the time for the completion is assumed to be only one year. Thirdly, the

[41] This variable is somewhat related to the export orientation of a province. Preliminary test shows that the correlation between this variable and the measure of export orientation (i.e. the ratio of exports to GDP) is above 0.6. Therefore, the export variable is not added in the model.

variable representing the IZs demand is the total number of enterprises, regardless of their ownership or industry. Here, the IZs demand should not be interpreted directly as the demand of enterprises to locate in the industrial zones. It is worth noting that, as presented in Chapter 2, not all types of economic activity are allowed to operate in these zones. In essence, the industrial zones have been targeted at attracting manufacturing activities, but this does not necessary mean that all manufacturing enterprises are located in the zones. It can be expected that the economic activities taking place in the zones could generate spillover effects to other economic activities located outside the zones but elsewhere in the province where these zones are located. In this way, they also generate beneficial effects for the entire provincial economy and, consequently, increase the competitive advantage of the province. This, in turn, influences the location choices of new enterprises. Therefore, the coefficient on the variable IZ should have a positive sign.

It is also necessary to examine the extent to which the ownership structure of the economy affects the locational behaviour of private sector enterprises. As introduced in Chapter 2, the number of SOEs has declined in recent decades. However, the state sector still maintains a large share of GDP, indicating its significant influence in the economy. The present study expects that the greater extent of the decrease in the state sector's dominance, the greater extent of the positive impact on the location choices of private sector enterprises. In this sense, the sign of the coefficient on the variable $SOECAP$, i.e. the share of SOEs in total physical capital stock of all enterprises located in the province, should be negative.

Similarly to the empirical analyses performed in Chapters 3 and 4, the average road density in neighbouring provinces is added to estimate the existence of cross-province spillover effect of transport infrastructure. This is aimed at examining the hypothesis that the location choices of enterprises may be influenced by not only the province's endowment of infrastructure but also the infrastructure endowment in its neighbouring provinces. The road spillover variable, $wROAD$, is calculated by using the two alternative spatial weights matrices, i.e. the one based on the contiguity relationship between provinces and the other one based on the contiguity and geographical-distance relationship between provinces, which have been described in Chapter 3.

5.3.3 Data

The firm-level data come from the Annual Enterprise Surveys conducted by the General Statistics Office of Vietnam. This chapter focuses on the location choices of new enterprises belonging to the private sector. If an enterprise, which is identified by its own tax code, appears for the first time in the merged firm-level dataset, it will be classified as new. For the purpose of empirical analysis, enterprises are categorised into the private sector if they were not 100 per cent owned by the central or local authorities. Concerning the firm-level data, enterprises with missing or zero values on key variables, namely, workforce, wage bill and fixed assets, are excluded from the sample.[42]

Although data are available from 2000, the present study examines the period 2002-2007. This is due to the calculation of some explanatory variables, such as the GDP growth and the industrial zone variables. After all, the total number of new private sector enterprises identified for the present study is 195,050. The geographical and sectoral distribution of these newly established enterprises is described below.

[42] Questionnaires concerning workforce and assets of enterprises include information on the value at the year-star and the value at the year-end. The year-end values are used as measures of the size of workforce and the stock of fixed assets (including long-term investments).

Disparities can be observed from the distribution of new private sector enterprises across size-based categories and sectors, as shown in Table 5.1. In the sectoral composition of the full sample, the leading sectors include manufacturing industries, trade activities and services. In the manufacturing sector, non-machinery enterprises outnumber machinery enterprises while, in the trade sector, wholesale strongly prevails retail. It can be observed from the distribution of new enterprises according to size-based categories that descriptive statistics for micro and large enterprise formation provide a significantly contrasting picture.

Table 5.1: Distribution of New Private Sector Enterprises by Workforce Sizes and Sectors

Workforce Sizes	No. of New Enterprises	Share in Total New Enterprises (%)	Sectors	No. of New Enterprises	Share in Total New Enterprises (%)
Micro	125,014	64.09	Non-Machinery Manufacturing	36,889	18.91
Small	50,475	25.88	Machinery Manufacturing	3,726	1.91
Medium	16,355	8.39	Wholesale Trade	41,294	21.17
Large	3,206	1.64	Retail Trade	24,216	12.42
			Services	19,595	10.05
			Other Sectors	69,330	35.54

Note:
1. Enterprises with missing or zero values on key variables were dropped from the full sample and sub-samples.
2. The year-end value is used as a measure of the size of workforce.
3. Machinery manufacturing sector includes machines, electrics, medical equipment, transport and other types of machines and equipment (2-digit industry codes: from 29 to 35). The 2-digit industry codes of non-machinery manufacturing are from 15 to 28, 36, and 37. The 2-digit industry codes of wholesale and retail are 51 and 52, respectively. Services include finance, insurance, property and other business services (2-digit industry codes: from 65 to 75).
4. In accordance with the definition stipulated in the Decree No. 90/2001/CP-NĐ of the Government of Vietnam, a micro enterprise has less than 10 employees; the workforce of a small enterprise ranges from 10 to 50; the workforce of a medium-sized enterprises ranges from more than 50 to less than 300; and a large enterprise has more than 300 employees.

To describe the spatial distribution of new private sector enterprises, two spatial levels are considered: region and province levels. Various publications of the General Statistics Office of Vietnam provide a standard classification of Vietnamese provinces into six regions, namely the Red River Delta, the Northern Mountain, the Central Coast, the Central Highlands, the Southeast and the Mekong River Delta. At the province level, there have been some categories of metropolis. The special and type-I metropolises, namely Ha Noi, Hai Phong, Da Nang, Ho Chi Minh City and Can Tho, are distributed vertically from the North to the South. Ha Noi and Hai Phong are located in the northern region, i.e. the Red River Delta. Da Nang is located in the central region, i.e. the Central Coast. Ho Chi Minh City and Can Tho are located in the southern regions, i.e. the Southeast and the Mekong River Delta, respectively. The Northern Mountain and the Central Highlands do not have any metropolis and, due to various factors ranging from geographical disadvantages and less-impressive economic achievements in comparison to other regions, can be considered the most backward regions of the country.

Table 5.2 shows that the most dynamic region is the Southeast which accounted for 37.06 per cent in the total number of new enterprises between 2002 and 2007, followed by the Red River Delta with 28.39 per cent. At the province level, Ho Chi Minh City had the largest share of new enterprises, followed by Ha Noi which accounted for 16.54 per cent. The shares of other metropolises in the total number of new enterprises are not as substantial as those of Ho Chi Minh City and Ha Noi, although their shares were higher than the average distribution rate of enterprises across non-metropolis provinces.

Table 5.2: Distribution of New Private Sector Enterprises by Regions and Metropolises

Regions	No. of New Enterprises	Share in Total New Enterprises (%)	Metropolises and the Rest	No. of New Enterprises	Share in Total New Enterprises (%)
Red River Delta (12 Provinces)	55,372	28.39	Ha Noi	32,258	16.54
			Hai Phong	5,750	2.95
Northern Mountain (13 Provinces)	10,298	5.28	-		
Central Coast (14 Provinces)	28,012	14.36	Da Nang	4,776	2.45
Central Highlands (4 Provinces)	5,570	2.86	-		
Southeast (6 Provinces)	72,295	37.06	Ho Chi Minh City	57,321	29.39
Mekong River Delta (11 Provinces)	23,503	12.05	Can Tho	3,556	1.82
			The Rest (56 Provinces)	91,389	46.85

Note: Enterprises with missing or zero values on key variables were dropped from the full sample and sub-samples which do not contain SOEs.

As can be seen in Table 5.3, there is no strong evidence indicating that micro, medium and large private sector enterprises locate differently across regions while it can only be seen from the small-enterprise category in which the share of the Red River Delta is slightly larger than that of the Southeast. In contrast, the sectoral figures provide more interesting insights into the spatially uneven distribution of enterprises although these figures generally confirm the Southeast, followed by the Red River Delta, as the most dynamic region. For all sectors of interest (i.e. apart from the group of unclassified industries), the share of the Southeast is larger than the aggregate figure reported in Table 5.2, and the most substantial increase can be observed in the service sector. With respect to the Red River Delta, its share in new enterprise formation decreases in retail and other manufacturing industries, compared with its aggregate figure, while increasing in machinery, wholesale and service sectors. For this region, the most substantial increase can be seen in the machinery sector. In contrast, Table 5.3 shows that the shares of the Northern Mountain, the Central Coast, the Central Highlands and the Mekong River Delta in total number of new enterprises belonging to the machinery and service sectors are smaller than their aggregate figures presented in Table 5.2. These figures indicate that machinery and service enterprises essentially cluster in the Southeast and the Red River Delta.

Table 5.3: Spatial Distribution of New Private Sector Enterprises by Workforce Size-Based Categories and Sectors (%)

	Red River Delta	Northern Mountain	Central Coast	Central Highlands	Southeast	Mekong River Delta
By Workforce Sizes:						
Micro	26.90	3.59	13.81	2.79	39.23	13.67
Small	32.53	8.79	15.94	3.04	29.66	10.05
Medium	27.47	7.39	14.30	2.98	41.11	6.77
Large	26.04	5.12	11.23	1.87	48.47	7.27
By Sectors:						
Non-Machinery	26.91	4.48	11.21	1.94	42.24	13.23
Machinery	38.46	2.90	7.57	0.46	46.03	4.59
Wholesale	36.60	3.50	9.68	2.13	41.37	6.72
Retail	18.99	3.53	18.24	2.25	39.10	17.88
Services	33.09	1.90	9.53	1.71	49.10	4.67
Other Sectors	25.70	8.46	19.20	4.44	27.16	15.05

Note: See Table 5.1

Table 5.4 presents descriptive statistics for count data of new private sector enterprises. The descriptive statistics show that, for the full sample, the mean number of new enterprises is around 532, while the standard deviation is over 1484, i.e. around 3 times the mean, indicating that the data are over-dispersed. Then, the assumption of the Poisson Regression model that the conditional variance equals the conditional mean does not appear to hold for the data of the full sample and several sub-samples of the present study.

Table 5.4: Descriptive Statistics for Count Data of New Private Sector Enterprises of the Full Sample and Sub-Samples including all Provinces

	Observations		Mean	Std. Dev.	Min	Max
	Total	Zeros				
Full Sample	366	0	532.9235	1484.536	23	15673
Micro	366	0	341.5683	1175.126	12	14322
Small	366	0	137.9098	310.7919	4	2993
Medium	366	1	44.68579	118.1582	0	1623
Large	366	57	8.759563	29.01945	0	452
Non-Machinery	366	0	100.7896	257.0632	2	2682
Machinery	366	109	10.18033	33.23936	0	358
Wholesale Trade	366	0	112.8251	436.8144	1	4768
Retail Trade	366	0	66.16393	205.6956	1	2783
Services	366	8	53.53825	250.6384	0	2844
Other Sectors	366	0	189.4262	361.7373	11	3669

Note: See Table 5.1

Another potential problem occurring with count data is an excess of zeroes. The zero-excess problem implies that there are two kinds of zeros thought to exist in the data, "true zeros" and "excess zeros", and the zero-inflated model attempts to account for this data problem by estimating two equations simultaneously, one for the count model and one for the excess zeros (e.g., Long, 1997; Cameron and Trivedi, 1998). However, what constitutes the incidence of excess zeros does not seem to be clarified in the literature. As can be seen in Table 5.4, the majority of sub-samples do not contain a zero value of the dependent variable. The zero-excess incidence may not be a serious econometric problem facing some sub-samples that contain zero observations. This is because the number of zero observations is rather small, in comparison to the number of non-zero observations, in most cases, enabling the application of the count regression model without zero inflation.

Finally, the panel dataset contains 61 provinces for 6 years. And, the total number of observations is 366. Descriptive statistics and correlation matrix of the explanatory variables mentioned in the previous section are presented in Tables 5.5 and 5.6. With respect to the sources of data, the road variable is calculated by using the data on road length provided by the Ministry of Transport and the data on area provided by the General Statistics Office of Vietnam; the data on the number of higher-education institutions are provided by the Ministry of Education and Training; the remaining variables are calculated by using the relevant data provided by the General Statistics Office of Vietnam.

Table 5.5: Descriptive Statistics of Explanatory Variables

Variable	Obs.	Mean	Std. Dev.	Min	Max
Log($ROAD$)	366	6.693	0.765	5.264	8.021
Log($w_{cont}ROAD$)	366	6.752	0.619	5.454	7.947
Log($w_{dist}ROAD$)	366	6.659	0.624	5.343	7.955
Log($POPULATION$)	366	13.95	0.526	12.54	15.68
$GDPGR$	366	11.21	3.955	-11.13	43.73
Log($WAGE$)	366	2.067	0.266	1.399	2.911
Log($WAGE$) ^{a}	366	1.947	0.301	0.995	2.714
Log($WAGE$) ^{b}	366	1.956	0.490	-0.036	3.144
Log($WAGE$) ^{c}	366	2.037	0.379	0.860	3.201
Log($WAGE$) ^{d}	366	1.840	0.324	1.007	2.810
Log($WAGE$) ^{e}	366	2.543	0.391	1.473	3.736
$NETMIG$	366	-0.731	5.195	-10.41	30.34
$HIGEDU$	366	3.535	7.840	0	57
$HERFINDAHL$	366	0.175	0.098	0.069	0.628
$LOCALISATION^{a}$	366	0.887	0.394	0.095	1.756
$LOCALISATION^{b}$	366	0.500	0.663	0	3.419
$LOCALISATION^{c}$	366	0.897	0.577	0.105	4.539
$LOCALISATION^{d}$	366	1.419	0.951	0.074	5.793
$LOCALISATION^{e}$	366	0.546	0.380	0.053	2.946
$SOECAP$	366	0.455	0.218	0.033	0.967
$SOECAP^{a}$	366	0.413	0.282	0	0.999
$SOECAP^{b}$	366	0.202	0.311	0	1
$SOECAP^{c}$	366	0.391	0.280	0	0.985
$SOECAP^{d}$	366	0.217	0.265	0	0.995
$SOECAP^{e}$	366	0.449	0.306	0	1
Log($ESIZE$)	366	4.031	0.511	2.790	5.209
IZ	366	0.320	0.541	0	3.359

Note:

1. All monetary variables are deflated by provincial GDP deflator.
2. The superscripts a, b, c, d and e denote the non-machinery manufacturing, machinery manufacturing, wholesale, retail and service sectors, respectively.

Table 5.6: Correlation Matrix of Explanatory Variables

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Log($ROAD$)	1.0000													
2	Log($w_{cont}ROAD$)	0.6874	1.0000												
3	Log($w_{dist}ROAD$)	0.7056	0.8877	1.0000											
5	Log($POPULATION$)	0.3691	0.4019	0.3760	1.0000										
6	$GDPGR$	0.0903	0.0221	-0.0147	-0.0713	1.0000									
7	Log($WAGE$)	0.0767	0.0727	-0.0008	0.2075	0.3322	1.0000								
8	Log($WAGE$)^{a}	0.0869	0.1483	0.0968	0.3116	0.1264	0.7477	1.0000							
9	Log($WAGE$)^{b}	0.2471	0.2117	0.2021	0.2421	0.1688	0.4575	0.3590	1.0000						
10	Log($WAGE$)^{c}	0.0237	0.0620	0.0663	0.2600	0.1113	0.6686	0.5430	0.3127	1.0000					
11	Log($WAGE$)^{d}	-0.1338	-0.1257	-0.1061	0.1197	0.1754	0.6824	0.5864	0.3743	0.6600	1.0000				
12	Log($WAGE$)^{e}	-0.1409	-0.1094	-0.0521	-0.0132	0.1191	0.4797	0.2756	0.1662	0.5130	0.5066	1.0000			
13	$NETMIG$	0.1334	-0.0616	-0.0054	0.1187	0.1505	0.2582	0.1200	0.2065	0.1630	0.2083	0.2041	1.0000		
14	$HIGEDU$	0.3850	0.2053	0.2756	0.5040	0.0134	0.2248	0.1595	0.2449	0.1694	0.1523	0.1514	0.5493	1.0000	
15	$HERFINDAHL$	-0.3284	-0.3187	-0.2652	-0.3401	0.1037	0.1072	-0.0179	-0.2487	0.0671	0.0874	0.2110	-0.0164	-0.2252	1.0000
16	$LOCALISATION^{a}$	0.4804	0.5617	0.4567	0.3581	0.0415	0.0443	0.2389	0.2010	0.1196	0.0113	-0.1398	-0.0205	0.0067	-0.3604
17	$LOCALISATION^{b}$	0.3881	0.3706	0.3226	0.3039	0.0842	0.1094	0.1225	0.4573	-0.0482	0.0399	-0.2230	0.1648	0.2803	-0.3875
18	$LOCALISATION^{c}$	0.0784	0.0365	0.0925	0.1623	-0.1551	-0.1685	-0.0856	-0.0844	-0.0797	-0.1956	-0.0191	0.0098	0.2181	-0.2415
19	$LOCALISATION^{d}$	-0.0879	-0.1290	-0.1131	-0.2299	-0.0941	-0.1909	-0.1587	-0.2379	-0.1045	-0.1589	0.0548	-0.1440	-0.1290	0.1763
20	$LOCALISATION^{e}$	0.1054	0.0244	0.1295	0.2571	-0.0611	0.1057	0.0799	0.0635	0.1107	0.0882	0.1820	0.3699	0.6832	-0.1411
21	$SOECAP$	-0.3277	-0.3519	-0.2881	-0.2234	-0.2601	-0.3195	-0.2158	-0.4011	-0.2981	-0.3102	-0.0958	-0.1055	-0.0935	0.1557
22	$SOECAP^{a}$	-0.3086	-0.2974	-0.2130	-0.2685	-0.2898	-0.3399	-0.2030	-0.3342	-0.2571	-0.2813	-0.0978	-0.0677	-0.0735	0.1637
23	$SOECAP^{b}$	0.0825	-0.0304	-0.0065	0.1508	-0.2072	-0.1362	-0.1268	-0.0588	-0.1120	-0.2176	-0.1441	-0.0508	-0.0065	-0.1940
24	$SOECAP^{c}$	-0.0668	-0.1543	-0.0988	-0.0033	-0.2087	-0.2302	-0.2231	-0.0933	-0.1512	-0.2097	0.0225	0.0732	0.0040	0.0772
25	$SOECAP^{d}$	-0.2094	-0.3503	-0.2631	-0.2002	-0.1859	-0.3557	-0.3544	-0.1781	-0.3627	-0.2319	-0.1444	0.0528	-0.0128	0.1988
26	$SOECAP^{e}$	-0.2124	-0.2131	-0.1576	-0.2203	-0.1550	-0.3156	-0.2448	-0.3357	-0.1950	-0.2369	0.0535	-0.1830	-0.2218	0.2425
27	Log($ESIZE$)	0.0619	-0.0260	-0.0524	0.1071	0.0590	0.0406	-0.0425	0.1752	-0.1913	-0.0100	-0.2623	0.2592	0.1496	-0.1193
28	IZ	0.1572	0.1592	0.0067	0.0775	0.1540	0.2428	0.1016	0.1878	0.1088	0.0969	-0.0312	0.1082	-0.0477	-0.2545

Note:
1. The sample contains 366 observations.
2. The superscripts a, b, c, d and e denote the non-machinery manufacturing, machinery manufacturing, wholesale, retail and service sectors, respectively.

Table 5.6 (cont.): Correlation Matrix of Explanatory Variables

		16	17	18	19	20	21	22	23	24	25	26	27	28
16	$LOCALISATION$ [a]	1.0000												
17	$LOCALISATION$ [b]	0.3159	1.0000											
18	$LOCALISATION$ [c]	-0.0442	0.0085	1.0000										
19	$LOCALISATION$ [d]	-0.2484	-0.2838	-0.0795	1.0000									
20	$LOCALISATION$ [e]	-0.2590	0.0647	0.3470	-0.0004	1.0000								
21	$SOECAP$	-0.3960	-0.3304	0.1890	0.2086	0.0744	1.0000							
22	$SOECAP$ [a]	-0.3208	-0.2869	0.0039	0.3056	0.0813	0.6850	1.0000						
23	$SOECAP$ [b]	-0.0210	0.0618	0.3517	0.0224	0.0321	0.2806	0.0906	1.0000					
24	$SOECAP$ [c]	-0.1137	-0.0987	0.2710	0.1262	0.1229	0.3614	0.2554	0.2826	1.0000				
25	$SOECAP$ [d]	-0.3334	-0.0639	-0.1621	0.2021	-0.0326	0.3725	0.3960	0.0984	0.4120	1.0000			
26	$SOECAP$ [e]	-0.2352	-0.3194	0.0478	0.2748	0.0198	0.5094	0.3205	0.1519	0.3785	0.3234	1.0000		
27	Log($ESIZE$)	0.0348	0.4398	-0.2914	-0.4093	-0.1309	-0.0458	-0.0809	0.0342	-0.0013	0.3284	-0.1540	1.0000	
28	IZ	0.3377	0.2220	-0.1267	-0.3514	-0.1806	-0.4026	-0.3212	0.0392	-0.0728	-0.1327	-0.2539	0.3909	1.0000

5.4 RESULTS

Results of the Negative Binomial estimation are presented in Tables 5.7, 5.8 and 5.9. In each table, regression (1) is estimated by assuming that all explanatory variables influence the locational behaviour of all categories of enterprises equally. Regressions (2), (3), (4) and (5) are specified for micro, small, medium and large enterprises, respectively. Finally, regressions (6), (7), (8), (9) and (10) are specified for enterprises belonging to the non-machinery, machinery, wholesale, retail and service sectors, respectively. The large test statistic of the log likelihood can be seen in all regressions, confirming the fit of the model. Wald χ^2 is statistically significant at the 1 per cent level, thus rejecting the null hypothesis that which all coefficients are simultaneous equal to zero. The following discussion is based on the results of the Negative Binomial Regression model; whereas, the results of the Poisson estimation are reported in Tables A3, A4, and A5 in Appendix A since the statistical significance of the Pearson χ^2 statistic, which falls below the standard threshold of 0.05, suggests that the Negative Binomial Regression model is more appropriate for the dataset employed.

5.4.1 Results of the Location-Choice Count Regression Model

Table 5.7: Results of the Negative Binomial Estimation of the Location-Choice Count Regression Model

	Full Sample	Workforce Size based Sub-Samples				Sector based Sub-Samples				
						Manufacturing		Trade		
		Micro	Small	Medium	Large	Non-Machinery	Machinery	Wholesale	Retail	Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(*ROAD*) $_{i,t-1}$	0.106*	0.091	0.090*	0.133***	0.217*	0.181***	0.206**	0.037	0.205**	0.202***
	(0.057)	(0.069)	(0.048)	(0.048)	(0.112)	(0.060)	(0.101)	(0.067)	(0.101)	(0.075)
Log(*POPULATION*) $_{i,t-1}$	0.619***	0.644***	0.635***	0.532***	0.442***	0.571***	0.797***	0.707***	0.509***	0.725***
	(0.054)	(0.072)	(0.051)	(0.057)	(0.126)	(0.073)	(0.125)	(0.073)	(0.110)	(0.085)
GDPGR $_{i,t-1}$	0.005	0.003	0.003	0.010**	0.000	0.006	0.005	0.014*	0.020**	-0.005
	(0.005)	(0.006)	(0.006)	(0.005)	(0.013)	(0.006)	(0.013)	(0.007)	(0.010)	(0.007)
Log(*WAGE*) $_{i,t-1}$	0.719***	0.951***	0.665***	0.113	0.294	0.246*	0.093	0.404***	0.557***	-0.004
	(0.116)	(0.144)	(0.124)	(0.123)	(0.192)	(0.127)	(0.109)	(0.126)	(0.148)	(0.101)
HIGEDU $_{i,t-1}$	0.045***	0.051***	0.033***	0.027***	0.031***	0.037***	0.031***	0.050***	0.051***	0.062***
	(0.004)	(0.005)	(0.004)	(0.003)	(0.006)	(0.004)	(0.007)	(0.005)	(0.007)	(0.007)
NETMIG $_{i,t-1}$	0.023***	0.016**	0.032***	0.024***	0.019**	0.030***	0.026***	0.026***	0.011	0.010
	(0.005)	(0.006)	(0.006)	(0.005)	(0.009)	(0.006)	(0.009)	(0.006)	(0.009)	(0.007)
HERFINDAHL $_{i,t-1}$	-0.769***	-0.886**	-0.884***	-1.282***	-1.784**	-1.286***	-3.360***	0.302	-0.110	-0.878*
	(0.261)	(0.343)	(0.262)	(0.323)	(0.840)	(0.376)	(0.996)	(0.504)	(0.520)	(0.481)
LOCALISATION $_{i,t-1}$						0.515***	0.300***	0.550***	0.289***	0.278**
						(0.092)	(0.089)	(0.093)	(0.056)	(0.139)
SOECAP $_{i,t-1}$	-0.262**	-0.129	-0.405***	-0.547***	-0.862***	-0.449***	0.072	-0.412***	-1.282***	-0.238*
	(0.124)	(0.153)	(0.125)	(0.128)	(0.254)	(0.124)	(0.160)	(0.155)	(0.227)	(0.126)
Log(*ESIZE*) $_{i,t-1}$	0.148**	0.003	0.158*	0.739***	1.071***	0.122	1.221***	0.508***	0.484***	0.465***
	(0.074)	(0.087)	(0.083)	(0.084)	(0.145)	(0.090)	(0.159)	(0.097)	(0.139)	(0.113)
IZ $_{i,t-1}$	0.111**	0.105	0.163***	0.006	0.033	0.201***	0.056	0.133*	0.156**	0.207***
	(0.056)	(0.069)	(0.055)	(0.054)	(0.101)	(0.059)	(0.108)	(0.076)	(0.075)	(0.067)
Constant	-5.031***	-5.979***	-5.844***	-7.161***	-9.759***	-5.633***	-15.476***	-9.474***	-7.613***	-10.723***
	(0.840)	(1.133)	(0.817)	(0.936)	(2.161)	(1.259)	(2.180)	(1.205)	(1.741)	(1.426)

Log likelihood	-2187.709	-2046.069	-1746.176	-1305.955	-829.113	-1618.332	-766.497	-1592.012	-1561.715	-1229.707
Wald χ^2 [p-value]	2009.148 [0.000]	1697.592 [0.000]	2029.851 [0.000]	3293.871 [0.000]	1325.869 [0.000]	2478.227 [0.000]	1599.696 [0.000]	2152.713 [0.000]	1006.841 [0.000]	2425.091 [0.000]
Observations	366	366	366	366	366	366	366	366	366	366
No. of Provinces	61	61	61	61	61	61	61	61	61	61

Note:

1. The dependent variable is the number of new enterprises belonging to the private sector.
2. All explanatory variables are lagged for one year. For the estimations of the full sample and workforce size based sub-samples, all explanatory variables are calculated at province level. For the sectoral estimations, the variables $WAGE$, $LOCALISATION$ and $SOECAP$ are calculated at the province-sector level while the rest are calculated at the province level.
3. Year and region dummies are included in all regressions.
4. Robust standard errors are in parentheses.
5. * p<0.100; ** p<0.050; *** p<0.010

Performance of Transport Infrastructure Variable

The improvement of road network generally attracts the location choices of enterprises. For the full sample, the magnitude of the coefficient on the variable $ROAD$, 0.106, indicates that for every 1 unit increase in the measure of road density, the log count of new private sector enterprises is expected to increase by 0.106 units, given that other explanatory variables are held constant in the model. However, the coefficient is only statistically significant at the 10 per cent level. This should indicate that the consideration on the importance of roads could vary across the location decision makers. The results obtained for different categories of enterprises suggest the complex nature of enterprises' transport infrastructure demand.

The statistical significance level of the coefficient on the variable $ROAD$ varies across the estimations using workforce-size-based sub-samples. Indeed, micro enterprises can be differentiated from other enterprises in regards to their input and output markets. Specifically, the markets of the former can be much smaller than those of the latter. This in turn may lead to the differences in their transport infrastructure demand. Consequently, the coefficient on the variable $ROAD$ is insignificant in the regressions specified for micro enterprises while the significantly positive results are obtained for small, medium and large enterprises. For the small, medium and large enterprises, the magnitude of the coefficient ranges from 0.090 to 0.217, which suggests that a 1 unit increase in the measure of road density leads to an expected increase in the log count of new private sector enterprises by 0.090, or 0.133, or 0.217 units, depending on the category of enterprises concerned and providing that other explanatory variables in the model are held constant. Given that transport infrastructure improvement is assumed to be positively associated with an increased access to the local market, competition could also be intensified by an improvement of market access and this might require enterprises to consider both their transport infrastructure demand and their own competitive capacity when choosing to locate in provinces with denser road networks. This could provide an additional explanation for the small magnitude of the coefficient, i.e. 0.09, estimated for micro enterprises whose capacity to compete with larger enterprises is supposed to be much weaker.

By the same token, the locational effect of the variable $ROAD$ varies across economic sectors. The results show that road infrastructure is an important factor determining the location choices of enterprises in the manufacturing, retail and service sectors but not in the wholesale sector. One explanation for the statistically insignificant result can be that the wholesale trading activities spread widely across provinces and, thus, the improvement of road network in a particular province alone could not meet the transport infrastructure demand of enterprises operating in this sector. On the contrary, the retailers could focus on serving the local market and, hence, the improved local road network, which can also be a proxy for an improvement of accessibility to local market, is more important for them in comparison to the wholesale traders. For the manufacturing, the retail and the service sectors, the magnitude of the coefficient is at around 0.2, which can be interpreted as that for an increase in the measure of road density by 1 unit, the log count of new private sector enterprises is expected to increase by around 0.2 units while holding other explanatory variables constant in the model.

Performance of Local Demand Variables

As far as the local market size measured by population is concerned, the coefficient on the variable $POPULATION$ is positive and statistically significant at the 1 per cent level in all cases. The magnitude of the coefficient ranges from 0.442 to 0.725, depending on the category of enterprises concerned. The modest magnitude suggests that a 1 unit increase in the measure of local market size leads to an expected increase in the log count of new private sector enterprises by 0.442 units, given that other explanatory variables are held constant in the model. The finding generally indicates that the size of local market is likely to exert a strong influence on the locational behaviour of enterprises. Given that this variable might capture the potential size of local labour pool and, under the assumption that the age structure was not significantly different between provinces, the result also suggests that the province with a larger population would have an additional advantage over those having smaller pools of inhabitants in attracting enterprises' location choices.

In contrast to the result for population, the growth rate of provincial GDP does not appear to be an influential factor shaping the location choices of enterprises in most cases. The coefficient on the variable $GDPGR$ only turns out to be statistically significant in the regression specified for medium enterprises and the trade sectors. It could be due to differences among enterprises in regard to the importance they place on the growth of demand in the local market.

Performance of Local Labour Market Variables

For the positive coefficient on the variable $WAGE$, it could be explained as follows. In theory, skilled employees could demand a higher wage level, in comparison to unskilled employees. Also, provinces with a higher degree of industrial agglomeration appear to have a higher wage level compared to that of less-concentrated provinces. In addition, according to Puga (2008) enterprises would not relocate to lower wage areas in response to an increase in the local wage level since they do not want to leave the current benefits of agglomeration, such as the proximity to suppliers of intermediate goods and to customers. Nevertheless, if the wages paid to skilled employees which are supposed to be already higher than those paid to unskilled employees increase further, then the coefficient on the variable $WAGE$ could turn out to be negative. In general, the positive sign of the coefficient on the variable $WAGE$ supports the assertion that, on average, enterprises could put much more emphasis on the quality of labour and business environment rather than on the labour costs when making their location choices. For the full sample, the magnitude of the coefficient, 0.719, suggests that for every 1 unit increase in the measure of the average wage level, the log count of new private sector enterprises is expected to increase by 0.719 units, given that other explanatory variables are held constant in the model.

Nevertheless, the coefficient on the variable $WAGE$ estimated for medium enterprises, large enterprises, and enterprises specialised in the machinery manufacturing and service sectors turns out to be statistically insignificant. Recall that, for the sectoral estimations, the average wage level is calculated at the province-sector level and therefore could capture the difference between sectors pertaining to their labour quality. Descriptive statistics presented in Table 5.5 show that the average wage level in the service sector is higher than that of the entire economy and those of other sectors. It can be expected that higher wages in the service sector, in comparison to other sectors, are to compensate employees for their own investment in education. Additionally to the latter argument, given that the value added of the service sector is higher than those of other sectors, on average, the labour quality of the former should be higher than that of the latter. Yet, the problem still exists, stemming from the data on average wages per employee, which do not capture the difference between enterprises in regard to the skill composition of their workforce, thus leaving the locational effect of the wage level in the machinery manufacturing and service sectors inconclusive.

Enterprises tend to establish in the province with a higher net migration rate. The coefficient on the variable $NETMIG$ is positive and statistically significant in most cases, except for the retail and service sectors. Accordingly, the variations in regard to the locational effect of the net migration rate can only be detected when taking into account the industry where the enterprises are specialised. For the full sample, the magnitude of the coefficient is 0.023, which indicates that for every 1 unit increase in the measure of net migration, the log count of new private sector enterprises is expected to increase by 0.023 units while holding other explanatory variables constant in the model. One explanation for the insignificant impact of the net migration rate on the location choices of retailers and service enterprises could be that these enterprises place a stronger emphasis on the local labour market. On the contrary, the significant results could suggest that the labour markets in which enterprises seek for employees are not necessarily local only. Given that the higher net migration rate could also imply an increase in the size of local market, and under the assumption that people migrated to provinces promising a higher wage level, which in turn implies a higher living standard, compared with that of their origin provinces, immigrants might not be the targeted customers of the retail and service sectors due to, on average and possibly in the first year of settlement, their lower purchasing power compared with that of the existing local inhabitants.

The coefficient on the variable $HIGEDU$ is positive and statistically significant at the 1 per cent level in all cases. The magnitude of the coefficient ranges from 0.027 to 0.062 across the full sample and sub-samples. The largest magnitude can be seen in the case of service enterprises, which suggests that a 1 unit increase in the measure of the capacity of qualified labour supply leads to an expected increase in the log count of new private sector enterprises by 0.062 units, given that other explanatory variables are held constant in the model. Pertaining to the hypothesis that different types of infrastructure could have different development outcomes, the locational effect of transport infrastructure is somewhat weaker, in terms of statistical significance, than that of higher-education infrastructure. This may indicate that the availability of higher-education institutions, which is used as a proxy for the capacity of qualified labour supply, could become more important over time, in comparison to physical infrastructure proxied by the road network, as a complementary factor of production.

Performance of Agglomeration Variables

Recall that a higher Herfindahl index implies a higher employment concentration by a single industry and, hence, lower industrial diversity. Result for the Herfindahl index indicates that enterprises, except for enterprises specialised in the trade sectors, prefer a diverse industrial environment. The statistically significant and negative coefficient on the variable $HERFINDAHL$ is evidence of a positive association between more industrial diversity and more industrial activity concentration. Put differently, the greater the industrial diversity in a province, the greater the attractiveness of that province to new enterprises which are seeking for an area to locate their factories. This result implies that new enterprises might need to rely on a diverse set of intermediate inputs from various industries. And, if a province concentrates too much on an industry, there are negative externalities that restrict its ability to attract new establishments. From another perspective, the importance of industrial diversity for new enterprises is consistent with the "nursery city" argument of Duranton and Puga (2001), which posits that enterprises are first established in more diverse areas where offer them a greater variety of local production knowledge and, consequently, help them to learn about their best production technology (Holl, 2004). For the full sample, the magnitude of the coefficient is -0.769, which indicates that for every 1 unit increase in the Herfindahl index measuring the lack of industrial diversity, the log count of new private sector enterprises is expected to decrease by -0.769 units while holding other explanatory variables in the model constant.

The variations in regard to the importance of industrial diversity for the location choices of enterprises can only be detected by taking into account the industry where the enterprises are specialised. In particular, the estimations using sectoral sub-samples reveal that the locational effect of industrial diversity depends on the sector concerned. In particular, the coefficient on the variable *HERFINDAHL* turns out to be statistically insignificant in the regressions specified for the wholesale and retail sectors. This indicates that the location choices of these enterprises are not affected by whether the industrial structure of the local economy is diverse or not. One explanation could be that, by nature, new trade enterprises do not involve in production activity and do not place a great importance on the diversity in the supply of intermediate inputs. Instead, they should focus on buying and selling activities which do not appear to be limited by provincial boundaries. In addition, some trade enterprises might not serve a wide range of industries and hence places a less emphasis on the industrial diversity of the local economy. Due to the lack of information on the number of industries that trade enterprises are expected to serve, the locational effect of industrial diversity is left inconclusive for the wholesale and retail sectors.

Results obtained for the variable *LOCALISATION* indicate that the higher the concentration of a sector in a particular province, the greater the attractiveness of that province to new establishments in that sector.[43] The results suggest that enterprises could expect the externalities, as presented in the literature review, associated with the concentration of same-sector economic activities. Moreover, the concentration of a particular industry in a particular province can reinforce the assessment that that industrial sector is feasible within

[43] For the full sample and workforce size-based sub-samples, due to the lack of information on the specialised industry of enterprises, the empirical analysis does not consider the locational effect of localisation economies and hence focus on the effect of urbanisation economies.

that province. The results for the intra-industry agglomeration variable might also indicate that the industrial location decision is path-dependent.

The positive agglomeration effect stemming from same-sector specialisation can be observed in all sectors. The magnitude of the coefficient on the variable $LOCALISATION$ ranges from 0.278 to 0.515, which suggests that a 1 unit increase in the measure of same-industry concentration leads to an increase in the log count of new private sector enterprises by the expected units ranging from 0.278 to 0.515, depending on the sector concerned, given that other explanatory variables are held constant in the model. However, in the regression specified for the service sector the coefficient is only statistically significant at the 5 per cent level, which is lower than those of other sectors. This could be influenced by the relatively strong concentration of the service sector, which leads to a higher level of competition and hence less possibility for newcomers to find market niches in the sector.

The variable $ESIZE$ has the expected positive sign. For the full sample, the magnitude of the coefficient, 0.148, suggests that for every 1 unit increase in the measure of the average size of enterprises located in a particular province regardless of their ownership and industry, the log count of new private sector enterprises is expected to increase by 0.148 units, while holding other explanatory variables in the model constant. The coefficient on the variable $ESIZE$ is statistically significant in all cases, expect for the case of micro enterprises. This indicates that in provinces with a higher level of entrepreneurship development, proxied by a larger average workforce size of enterprises, new micro enterprises might be unable to either compete with the existing larger enterprises or to overcome business-costs related entry barriers.

Performance of Public Policy Variables

The coefficient on the variable $SOECAP$ is negative and statistically significant in most cases, except for the cases of micro enterprises and machinery-manufacturing enterprises. For the full sample, the magnitude of the coefficient, -0.194, indicates that for every 1 unit increase in the proportion of physical capital stock owned by SOEs in the total stock of physical capital, the log count of new private sector enterprises is expected to decrease by -0.194 units, given that other explanatory variables in the model are held constant. The negative relationship between the share of SOEs in total physical capital stock and the location choices of enterprises belonging to the private sector suggests that the scaled-down restructuring of the state sector, apart from other local attributes, have a significant explanatory power for the explosion of the private entrepreneurship.

Another policy variable is the one capturing the provision of industrial zones. At the aggregate level, the provision of industrial zones within a province shows no role in determining the choices of enterprises to locate in that province. In particular, the magnitude of the coefficient, 0.111, indicates that for every 1 unit increase in the measure of industrial zone supply, the log count of new private sector enterprises is expected to increase by 0.111 units, while holding other explanatory variables in the model constant. Result for the variable IZ indeed suggests that the locational effect of industrial zones varies dramatically according to the enterprise-specific characteristics.

In the estimations of workforce size-based sub-samples, the coefficient on the variable IZ is only statistically significant, at the 5 per cent level, in the case of small enterprises. The results for the workforce size-based sub-samples could not be properly explained since the industry in which enterprises are specialised is unknown. From the sectoral perspective, non-machinery, trade and service enterprises are strongly influenced by the province's endowment of industrial zones when making their location choices. This could be that these enterprises tend to supply their services to the enterprises located in the zones. Alternatively, there might be some kind of spillover effects from the concentration of industrial activity in the zones that contributes to the vibrant business environment of the entire province which in turn generates another advantage for the province to compete for new location choices.

The insignificant result for the variable IZ obtained from estimating the regression specified for the machinery manufacturing sector could be explained by the fact that the technology-level as well as the export-orientation related heterogeneities among these zones and machinery manufacturing enterprises could not be captured by the variable IZ and the count number of new enterprises, respectively. Therefore, the importance of the availability of industrial zones in shaping the location choices of machinery manufacturing enterprises is inconclusive.

5.4.2 Results of the Location-Choice Count Regression Model with Road Spillover Variable

The estimation of the model with road spillover variable reveals mixed findings concerning the existence of cross-province road spillover effect. In particular, the performance of the road spillover variable is dependent on the spatial weights matrix employed. As can be seen in Table 5.8, the coefficient on the road spillover variable calculated by using the contiguity-based spatial weights matrix, $w_{cont}ROAD$, is statistically insignificant in all cases, indicating that there is no evidence to empirically confirm the existence of cross-province road spillovers.

Results presented in Table 5.9 show some evidence of positive cross-province road spillovers. The coefficient on the road spillover variable calculated by using the contiguity- and distance-based spatial weights matrix, $w_{dist}ROAD$, is positive and statistically significant in some cases. In fact, that positive spillover effect varies dramatically across the categories of enterprises examined.

Furthermore, the introduction of the road spillover variable into the regression affects the performance of the road variable. In particular, apart from the retail and service sub-samples, the estimation for the full sample and other sub-samples shows that both the magnitude and the statistical significance of the coefficient on the variable $ROAD$ almost reduce along with the inclusion of the variable $w_{dist}ROAD$. One explanation for the retail and service sectors could be that the majority of enterprises specialised in these sectors focus on serving the local market rather than the neighbouring markets and/or do not place a great importance on the neighbouring markets as a source of inputs, and they might therefore not be concern about how the neighbouring road network has been improved.

Table 5.8: Results of the Negative Binomial Estimation of the Location-Choice Count Regression Model with w_{cont} Road Spillover Variable

	Full Sample	Workforce Size based Sub-Samples				Sector based Sub-Samples				
						Manufacturing		Trade		
		Micro	Small	Medium	Large	Non-Machinery	Machinery	Wholesale	Retail	Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(*ROAD*) $_{i,t-1}$	0.105*	0.087	0.091*	0.131***	0.215*	0.181***	0.206**	0.037	0.200**	0.203***
	(0.057)	(0.068)	(0.049)	(0.047)	(0.112)	(0.061)	(0.101)	(0.067)	(0.100)	(0.075)
Log(w_{cont}*ROAD*) $_{i,t-1}$	0.044	0.104	-0.097	0.096	0.106	-0.054	0.024	0.002	0.126	-0.089
	(0.077)	(0.094)	(0.076)	(0.090)	(0.150)	(0.104)	(0.175)	(0.113)	(0.142)	(0.133)
Log(*POPULATION*) $_{i,t-1}$	0.613***	0.628***	0.649***	0.517***	0.427***	0.578***	0.793***	0.706***	0.492***	0.737***
	(0.057)	(0.075)	(0.052)	(0.059)	(0.126)	(0.074)	(0.126)	(0.076)	(0.115)	(0.088)
GDPGR $_{i,t-1}$	0.005	0.004	0.003	0.011**	0.001	0.005	0.005	0.014*	0.021*	-0.006
	(0.005)	(0.006)	(0.006)	(0.005)	(0.013)	(0.006)	(0.013)	(0.007)	(0.010)	(0.007)
Log(*WAGE*) $_{i,t-1}$	0.727***	0.969***	0.645***	0.134	0.320*	0.243*	0.095	0.404***	0.581***	-0.006
	(0.118)	(0.147)	(0.128)	(0.126)	(0.193)	(0.128)	(0.112)	(0.128)	(0.150)	(0.101)
HIGEDU $_{i,t-1}$	0.045***	0.051***	0.032***	0.027***	0.031***	0.037***	0.031***	0.050***	0.052***	0.062***
	(0.004)	(0.005)	(0.004)	(0.003)	(0.006)	(0.004)	(0.007)	(0.005)	(0.007)	(0.007)
NETMIG $_{i,t-1}$	0.023***	0.016**	0.032***	0.024***	0.020**	0.030***	0.026***	0.026***	0.011	0.010
	(0.006)	(0.006)	(0.006)	(0.005)	(0.009)	(0.006)	(0.009)	(0.006)	(0.009)	(0.007)
HERFINDAHL $_{i,t-1}$	-0.720***	-0.772**	-0.995***	-1.175***	-1.663*	-1.332***	-3.324***	0.304	0.031	-0.992*
	(0.275)	(0.360)	(0.279)	(0.336)	(0.856)	(0.392)	(1.042)	(0.513)	(0.548)	(0.523)
LOCALISATION $_{i,t-1}$						0.529***	0.301***	0.550***	0.287***	0.266*
						(0.096)	(0.089)	(0.092)	(0.056)	(0.141)
SOECAP $_{i,t-1}$	-0.260**	-0.126	-0.409***	-0.546***	-0.861***	-0.445***	0.077	-0.412**	-1.264***	-0.232*
	(0.124)	(0.152)	(0.126)	(0.127)	(0.254)	(0.124)	(0.166)	(0.154)	(0.227)	(0.127)
Log(*ESIZE*) $_{i,t-1}$	0.147**	0.003	0.160*	0.737***	1.067***	0.120	1.218***	0.508***	0.477***	0.461***
	(0.074)	(0.087)	(0.083)	(0.083)	(0.145)	(0.090)	(0.161)	(0.097)	(0.139)	(0.113)
IZ $_{i,t-1}$	0.110**	0.102	0.165**	0.003	0.033	0.202***	0.055	0.133*	0.154*	0.208***
	(0.055)	(0.068)	(0.056)	(0.054)	(0.100)	(0.059)	(0.108)	(0.077)	(0.073)	(0.067)
Constant	-5.276***	-6.545***	-5.293***	-7.679***	-10.355***	-5.329***	-15.600***	-9.483***	-8.280***	-10.211***
	(0.929)	(1.232)	(0.917)	(1.022)	(2.338)	(1.380)	(2.424)	(1.332)	(1.846)	(1.608)

Log likelihood	-2187.578	-2045.584	-1745.575	-1305.395	-828.895	-1618.211	-766.488	-1592.012	-1561.366	-1229.474
Wald χ^2 [*p*-value]	2008.593 [0.000]	1701.486 [0.000]	2072.928 [0.000]	3254.734 [0.000]	1347.009 [0.000]	2483.220 [0.000]	1601.490 [0.000]	2153.807 [0.000]	999.129 [0.000]	2489.834 [0.000]
Observations	366	366	366	366	366	366	366	366	366	366
No. of Provinces	61	61	61	61	61	61	61	61	61	61

Note: See Table 5.7

Table 5.9: Results of the Negative Binomial Estimation of the Location-Choice Count Regression Model with w_{dist} Road Spillover Variable

	Full Sample	Workforce Size based Sub-Samples				Sector based Sub-Samples				
						Manufacturing		Trade		Services
		Micro	Small	Medium	Large	Non-Machinery	Machinery	Wholesale	Retail	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(*ROAD*) $_{i,t-1}$	0.083	0.070	0.071	0.095**	0.182	0.153***	0.149	0.056	0.171*	0.213***
	(0.052)	(0.064)	(0.047)	(0.047)	(0.117)	(0.059)	(0.102)	(0.070)	(0.098)	(0.074)
Log(w_{dist}*ROAD*) $_{i,t-1}$	0.149**	0.124	0.133**	0.224***	0.182	0.175**	0.261*	-0.120	0.162	-0.056
	(0.060)	(0.075)	(0.058)	(0.062)	(0.116)	(0.080)	(0.145)	(0.089)	(0.110)	(0.098)
Log(*POPULATION*) $_{i,t-1}$	0.603***	0.631***	0.619***	0.498***	0.416***	0.549***	0.760***	0.720***	0.498***	0.733***
	(0.055)	(0.074)	(0.052)	(0.058)	(0.127)	(0.075)	(0.129)	(0.074)	(0.112)	(0.086)
GDPGR $_{i,t-1}$	0.006	0.004	0.004	0.012**	0.002	0.007	0.010*	0.012	0.022**	-0.006
	(0.005)	(0.006)	(0.005)	(0.005)	(0.013)	(0.006)	(0.014)	(0.007)	(0.010)	(0.007)
Log(*WAGE*) $_{i,t-1}$	0.783***	1.001***	0.726***	0.215*	0.378*	0.283**	0.101	0.401**	0.579***	-0.000
	(0.122)	(0.153)	(0.127)	(0.128)	(0.200)	(0.126)	(0.110)	(0.126)	(0.149)	(0.100)
HIGEDU $_{i,t-1}$	0.044***	0.051***	0.032***	0.026***	0.031***	0.037***	0.031***	0.050***	0.051***	0.062***
	(0.004)	(0.005)	(0.004)	(0.003)	(0.006)	(0.004)	(0.007)	(0.005)	(0.007)	(0.007)
NETMIG $_{i,t-1}$	0.021***	0.015**	0.031***	0.023***	0.018**	0.029***	0.024***	0.026***	0.010	0.011
	(0.005)	(0.006)	(0.006)	(0.005)	(0.009)	(0.006)	(0.009)	(0.006)	(0.009)	(0.007)
HERFINDAHL $_{i,t-1}$	-0.638**	-0.770**	-0.781***	-1.095***	-1.599*	-1.156***	-2.912***	0.157	0.056	-0.943*
	(0.263)	(0.347)	(0.259)	(0.321)	(0.845)	(0.377)	(1.018)	(0.503)	(0.536)	(0.508)
LOCALISATION $_{i,t-1}$						0.473***	0.323***	0.550***	0.296***	0.277**
						(0.096)	(0.088)	(0.093)	(0.056)	(0.140)
SOECAP $_{i,t-1}$	-0.255**	-0.125	-0.394***	-0.542***	-0.863***	-0.474***	0.135	-0.403**	-1.286***	-0.232*
	(0.122)	(0.151)	(0.122)	(0.122)	(0.254)	(0.124)	(0.167)	(0.154)	(0.223)	(0.128)
Log(*ESIZE*) $_{i,t-1}$	0.136*	-0.006	0.148*	0.718***	1.046***	0.118	1.181***	0.517***	0.483***	0.468***
	(0.073)	(0.087)	(0.082)	(0.083)	(0.144)	(0.090)	(0.160)	(0.096)	(0.140)	(0.113)
IZ $_{i,t-1}$	0.119**	0.113*	0.167**	0.016	0.051	0.209***	0.075	0.121	0.169**	0.202***
	(0.054)	(0.067)	(0.054)	(0.052)	(0.101)	(0.057)	(0.108)	(0.075)	(0.073)	(0.067)
Constant	-5.823***	-6.637***	-6.548***	-8.186***	-10.555***	-6.422***	-16.459***	-8.933***	-8.454***	-10.513***
	(0.918)	(1.225)	(0.836)	(0.957)	(2.211)	(1.294)	(2.099)	(1.269)	(1.817)	(1.495)

Log likelihood	-2185.268	-2044.957	-1744.306	-1301.023	-828.097	-1616.205	-764.557	-1591.238	-1560.707	-1229.547
Wald χ^2 [p-value]	2051.962 [0.000]	1714.660 [0.000]	2125.590 [0.000]	3453.447 [0.000]	1347.768 [0.000]	2601.657 [0.000]	1686.119 [0.000]	2141.892 [0.000]	1003.792 [0.000]	2437.527 [0.000]
Observations	366	366	366	366	366	366	366	366	366	366
No. of Provinces	61	61	61	61	61	61	61	61	61	61

Note: See Table 5.7

5.5 SUMMARY AND POLICY DISCUSSION

The empirical analysis presented in Chapter 5 provides some main findings, as summarised below.

The improvement of the road network, which is assumed to be associated with the improvement of local market access, increases the attractiveness of the province to private sector enterprises. The empirical analysis also shows that the evidence of the cross-province spillovers of transport infrastructure varies dramatically according to the categories of enterprises examined and turns out to be obviously dependent on the spatial weights matrix used to measure the neighbouring road network. Other local attributes that are positively associated with the number of private sector enterprises choosing to locate in the province include the size of local market, the capacity of labour supply, especially qualified labour supply, the development of local entrepreneurship, the economies of same-industry agglomeration, and the industrial diversity of the provincial economy. Finally, the scaled-down restructuring of SOEs relative to the whole economy and the provision of industrial zones have a role to play in explaining why some particular provinces are chosen as the locations of private sector establishments.

The results for the estimations of sub-samples based on workforce size and industry confirm the differences in locational behaviour according to a number of enterprise-specific characteristics. With respect to transport infrastructure, the results for different categories of enterprises suggest the complex nature of enterprises' transport infrastructure demand. The locational effects that hold across different sub-samples are those of the local market size and the higher-education infrastructure.

Several policy implications can be drawn from the empirical analysis, as presented below.

Firstly, policies aimed at increasing the road density, i.e. the length of roads per 1,000 km^2 of the provincial area, to raise the comparative advantage of a particular province in attracting the location choices of newly-established private sector enterprises is generally supported by empirical evidence found in the present study. However, the finding on the performance of the road variable indicates that the impact of transport infrastructure varies across different sub-samples as long as enterprise-specific characteristics are taken into account. In contrast, there is robust evidence indicating that the endowment of higher-education institutions is a local advantage in competing for private sector location choices. In essence, this finding lends a support to the importance of local capacity of qualified labour supply.

Secondly, the results for the Herfindahl index imply that the lack of industrial diversity would be detrimental to the attractiveness of the province to the industrial location choices. However, creating a diverse industrial structure does not suggest an increase in the number of industries; instead, it implies a suggestion for an assessment on whether resources, such as labour, have been allocated efficiently across industries so as to avoid the fact that resources are not over-concentrated in a single industry. As a practical issue, it is important to note that the Herfindahl index could not control the possibility that some provinces might have been the location of the industries which were not at their competitive advantages. In this case, resources might have been used inefficiently.

Thirdly, although intra-industry agglomeration is found to be an advantage of the province in attracting the location choices of private sector enterprises, the nationwide dimension of policies aimed at promoting the intra-industry agglomeration should be taken into account. It is worth recalling the "virtuous circles of self-reinforcing development", suggested by

Venables (2005) in the context of developing countries, which could be created in some geographical areas by agglomeration forces and, consequently, would leave other areas lag further behind. For economic centres, an increased concentration of economic activities does not always promise beneficial effects and, especially, when the development of infrastructure and other amenities does not go hand in hand with that increased concentration, agglomeration economies will then be outweighed by such diseconomies as environmental pollution, traffic congestion, excess commuting, higher wages and high land prices. In this sense, it is important for the central government to provide a nationwide-level policy of specialisation with an attention to the efficiency of resource allocation throughout the country, the trade-offs between spatial disparities and intra-industry agglomeration economies, as well as the adverse effects of over-concentration in a few economic centres.

Fourthly, the empirical evidence suggests that, in general and for some specific sectors, the scaled-down restructuring of the SOE sector should be one important component of polices seeking to promote private sector development and, in particular, to accelerate the formation of new private sector enterprises, which in turn determines the development of the national economy as well as and the creation of job opportunities. The empirical results for the ownership structure variables obtained in Chapters 3 and 4 and in this chapter together support the fact that the economic reforms leading to the scaled-down restructuring of the SOE sector were not only beneficial for the explosion of the private sector but also for the growth of the whole economy over the sample period.

Finally, that the set of local attributes attracting the location choices of enterprises varies across different economic sectors suggests that the province should consider its own advantages as well as disadvantages when making policies aimed at attracting the location choices of specific categories of private sector enterprises. Additionally, it is important to emphasise that micro, small, medium and large enterprises would have different responses to local attributes when making their location choices. This in turn implies that policies aimed at promoting the explosion of private entrepreneurship should also take into account the fact that enterprises at different development stages, which can in part be captured by the average size of their workforce, would have different needs which in turn affect their location preferences. After all, policies capturing the heterogeneities among categories of enterprises would be more favourable than those treating all categories of enterprises identical.

CHAPTER 6

CONCLUSION

As stated in the introductory chapter, the contribution of this thesis is intended to be empirical only. The research is aimed to examining the impact of infrastructure on the three aspects of the Vietnamese economy, namely, economic growth, private sector employment and private sector location choices. Findings obtained in the thesis provide an insight into the role of infrastructure in Vietnam and add further empirical evidence into the on-going academic debate on the economic effects of infrastructure. Also, those findings have a variety of important policy implications, beyond implications for the public policy of infrastructure provision.

Chapter 6 begins with a summary of main findings. The next section presents policy implications. Limitations of the empirical studies and, then, recommendations for further research are presented in the final section.

6.1 MAIN FINDINGS

6.1.1 Findings on the Role of Infrastructure

The empirical evidence obtained in Chapter 3 suggests that transport infrastructure is an important engine of economic growth. This conclusion holds across various specifications of the growth regression. In Chapter 4, the empirical analysis shows that transport infrastructure is an important determinant of private sector employment. This impact generally is not sensitive to what other explanatory variables are included in the employment regression. Finally, the empirical analysis presented in Chapter 5 confirms that transport infrastructure is an influential factor shaping the location choices of private sector enterprises. In particular, an improvement in the road network enhances the attractiveness of a particular province to enterprises belonging to the private sector. However, the disaggregate analysis reveals that transport infrastructure does not appear to be of a statistical importance in explaining the location choices of several categorises of private sector enterprises. This in turn implies the complex nature of the private sector's transport infrastructure demand.

Taken together, these findings suggest that transport infrastructure had a role to play in the Vietnamese economy over the sample period. In particular, the empirical research shows that an improved transport infrastructure endowment is important for promoting economic growth, stimulating private sector employment and attracting the location choices of newly-established private sector enterprises. Returning to the research questions posed at the beginning of the thesis, it is now possible to state that an increased endowment of transport infrastructure is important for achieving economic outcomes. However, based on the findings obtained in the location choice analysis, it is necessary to emphasise that the important role of transport infrastructure should also be dependent on the case concerned.

Higher-education infrastructure is another category of infrastructure examined in Chapters 4 and 5. There is strong empirical evidence of the positive impact of higher-education infrastructure, which is assumed to be a proxy for the capacity of qualified labour supply, on private sector employment and location choices. Additionally, the statistical importance of higher-education infrastructure holds in all sectors considered in the location choice analysis, emphasising further the importance of the provision of higher-education infrastructure and, especially, the capacity of qualified labour supply.

I is worth emphasising that the infrastructure variables examined in the empirical research of the thesis are expressed in physical terms, not monetary terms. In particular, the empirical findings on the beneficial effects of infrastructure presented in the thesis do not imply any suggestion in relation to an increase in public spending on infrastructure. Instead, they should be interpreted from the perspective of an improved capacity of infrastructure.

6.1.2 Findings on the Cross-Province Transport Infrastructure Spillovers

The possibility that the provision of transport infrastructure in neighbouring provinces also affects economic performance of a particular province has been explicitly taken into account in the empirical analyses. The neighbour relationship is defined basing on the geographical contiguity. Alternative spatial weights matrices have also been tried to measure the road spillover variable.

In the growth analysis, there is no evidence supporting the hypothesis on the existence of cross-province spillovers from transport infrastructure. However, regardless of the variations in the statistical significance of the road spillover variable, there is some evidence suggesting the existence of a negative cross-province spillover effect on private sector employment. This could be driven by the fact that transport infrastructure improvement alters the comparative advantages across provinces, thereby drawing economic activities, in particular job opportunities, from elsewhere to the provinces where such the improvement takes place. On the contrary, the location choice analysis provides some evidence of positive cross-province transport infrastructure spillovers. However, that positive spillover effect is dependent on the spatial weights matrix used to calculate the road spillover variable and also varies dramatically according to the categories of enterprises examined.

6.1.3 Other Main Findings

The growth analysis provides evidence of convergence. With respect to the traditional growth elements, the growth of physical capital investment appears to be an important driver of economic growth while the relationship between the growth of human capital supply and the growth of GDP per capita is unable to be identified. Regarding other determinants of growth, the performance of inflation measured by the change in the logarithm of the consumer price index, the GDP sectoral transformation, and the ownership structure are of statistical importance. In particular, inflation turns out to be negatively associated with growth. A faster shift towards the economy which is less dependent on agriculture leads to a higher overall growth. Finally, a decrease in the proportion of SOEs in the total number of enterprises is positively associated with an increase in the overall growth.

In regard to the determinants of private sector employment, there is evidence of a positive relationship between private sector employment and production scale, as well as between the numbers of private sector employees and enterprises, but a negative relationship between private sector employment and productivity growth. With respect to the impact of the labour supply factors on private sector employment, the capacity of qualified labour supply and the degree of urbanisation have positive and statistically significant impacts while the statistical significance of the positive employment impacts of the high-school graduate growth and the net migration rate varies according to the model specifications. Finally, the proportion of SOE employment in total employment is found to be negatively associated with private sector employment.

Beyond infrastructure, other local attributes that attract the location choices of enterprises include the size of the local market, the capacity of qualified labour supply, the development of local entrepreneurship, the same-industry concentration of enterprises, the industrial diversity of the provincial economy, the scaled-down restructuring of the SOE sector, and the provision of industrial zones. The results for the estimations using sub-samples lead to the conclusion that there are notable differences in the locational behaviour according to a number of enterprise-specific characteristics. Finally, the locational effects that hold in all sub-samples are those of the local market size and higher-education infrastructure variables.

As stated early, the thesis does not attempt to empirically assess the role of SOEs in the Vietnamese economy, and the inclusion of the ownership structure variables is aimed at capturing one of the striking characteristics of the economy in transition of Vietnam. However, given that the focus of the thesis is on the role of infrastructure which has been essentially provided by the state sector, the performance of the ownership structure variable, i.e. the variable measuring the proportion of SOEs in physical capital stock, or in employment, or in the total number of enterprises, should be discussed further. As found in the growth analysis, a decrease in the proportion of SOEs in the total number of enterprises located in a particular province leads to an increase in the overall growth of that province. The second empirical analysis shows that a decreased proportion of SOE employment in total employment is positively associated with an increased private sector employment. Finally, there is evidence of the negative relationship between the proportion of physical capital stock owned by SOEs in total physical capital stock and the location choices of private sector enterprises. In general, these empirical findings suggest that the scaled-down restructuring of the SOE sector relative to the whole economy, apart from other province-specific characteristics, have a significant explanatory power for the overall growth performance and the evolution of the emerging private sector. This conclusion is indeed in line with the on-going economic reforms toward encouraging the development of the private sector in the Vietnamese economy that have been gradually implemented since the late 1980s. However, as also stated in the previous chapters, caution is required for the interpretation of these empirical findings due to the fact that Chapters 3 and 4 are unable to take into account the sectoral differences in the performance of explanatory variables, including the ownership structure variable. This is in part supported by the statistical variations in the performance of the ownership structure variable in the disaggregate location choice analysis using sectoral sub-samples.

6.2 POLICY IMPLICATIONS

6.2.1 Implications for Infrastructure Provision

The above-summarised empirical findings have important implications for transport infrastructure policy towards accelerating GDP growth, stimulating private sector employment and attracting the location choices of private sector enterprises. In particular, policies aimed at increasing the road length per 1,000 km^2 of the provincial area to achieve such economic outcomes as GDP growth, private sector employment and industrial location choices are empirically supported. However, as stated in each empirical chapter, the expectation on the positive effect of infrastructure is based on a number of assumptions which are essentially related to the quality of infrastructure. Therefore, it is important to examine the quality of infrastructure in a particular province as well as the heterogeneity in regard to the infrastructure quality among provinces while using the empirical findings presented in the thesis to suggest any infrastructure-related policies.

The firm-level data allows the test on whether the effect of infrastructure varies according to the categories of enterprises examined. In particular, there is strong evidence of a variance in the locational effect of transport infrastructure across sub-samples capturing specific characteristics of enterprises. This suggests that the use of infrastructure improvement as a policy instrument might not bring about the expected outcomes if the heterogeneities among the targeted beneficiaries of infrastructure provision were not taken into account.

The effects of transport infrastructure could be either underestimated or overestimated if the possibility of the spatial spillovers of such the effects is ignored, as presented in the literature reviews. In terms of policy implications, the following conclusions are based on the empirical finding obtained in Chapter 4 which shows that an improvement of transport infrastructure in the neighbouring provinces is detrimental to the employment gained in a particular province. First, policies made at the provincial level should also take into account the competitive pressure arising from the development of neighbouring transport networks. Second, beyond infrastructure-related policies, other targeted province-specific policies for the provinces with underdeveloped infrastructure systems are needed to avoid the possibility that these provinces would be left further behind. Third, due to the network characteristics of transport infrastructure, the central government should pay special attention to the provincial coordination of transport construction to minimise the adverse effect of the fact that, by altering investment patterns in transport infrastructure relative to those of the neighbouring provinces, each province has the ability to modify the size of its transport stock and hence achieve its economic outcomes at the expense of its neighbours.

With respect to higher-education infrastructure, the empirical findings suggest the importance of the provision of this category of infrastructure. However, it is worth recalling that the estimation of the effect of higher-education infrastructure is based on the assumption that a higher endowment of higher-education infrastructure is positively associated with a higher capacity of qualified labour supply. Therefore, policies related to higher education should also explicitly take into account the capacity of higher-education institutions in training and providing qualified labour.

6.2.2 Other Policy Implications

In addition to the policy implications for infrastructure provision, Chapter 3 also suggests other implications for growth-enhancing polices. In particular, policies to accelerate the process of structural transformation toward industrialisation can also promote the overall economic growth. With respect to the structural transformation in terms of ownership, the economic reforms aimed at encouraging an increased role of the private sector can accelerate the growth of the economy as a whole. Policies to reduce the volatility of the provincial economy to macroeconomic shocks, particularly the shocks associated with the CPI performance, may be critical for achieving the growth target.

In terms of policy implications, some conclusions can be drawn from chapter 4. It appears that the efficient way to stimulate private sector employment would be to expand the production process, encourage the private sector establishments, enhance the capacity of qualified labour supply, and also promote the ownership restructuring towards a decline in the size of the SOE sector relative to the whole economy. Given the negative relationship between private sector employment and productivity growth, which could suggest that a proportion of unskilled workers might be left unemployed and that some of them might join in the informal sector of the labour market, policies to improve the quality of the labour force would contribute to controlling for the issue of unemployment facing unskilled and semi-skilled workers who have been a dominant force of the large labour pool of Vietnam. Last but not least, although the degree of urbanisation is positively related to private sector employment, urbanisation-related policies need to take into account the phenomenon that the private sector economic activities should be concentrated in the urban centres and the fact that the congestion-related adverse effects arising from this concentration should be controlled for when encouraging the explosion of the private sector establishments.

In Chapter 5, in addition to the policy implications for infrastructure provision, there are other implications for the use of policy instruments. In particular, it would be necessary to improve other provincial characteristics, such as industrial diversity, intra-industry agglomeration, and industrial zones, in order to attract the location choices of private sector enterprises, as suggested in Chapter 5. Similarly to the employment issue, the scaled-down restructuring of the SOE sector should be one important component of the set of polices seeking to promote the private sector development. Given that the set of local attributes determining the private sector location choices varies across different categories of enterprises, each province should consider its own comparative advantages as well as disadvantages when making policies aimed at attracting the location choices. Finally, policies controlling for the specific characteristics and the specific needs of each category of enterprises would be more favourable than those treating all enterprises identical.

6.3 LIMITATIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

6.3.1 Limitations

The thesis is limited in several ways, mostly because of data limitations. In fact, it is acknowledged that the limitations of data always represent a constraint for empirical research, and this is especially true for research on developing countries due to their immature statistical system (Straub, 2008).

In the present study which focuses on the role of infrastructure, the most important limitation lies in the fact that the empirical analyses are unable to control the quality of infrastructure while examining its economic effects. Since the level of development has been very different between provinces, it appears to be unrealistic to assume that the quality of infrastructure is homogenous across these geographical areas. In addition, due to the lack of information on the composition of road infrastructure, the empirical research is unable to capture the fact that different types of infrastructure would have different contributions to achieving the economic outcomes examined in the thesis. Moreover, the thesis is unable to provide a general conclusion on the cross-province spillover effect of transport infrastructure. Apart from the data issues, the choice of spatial weights matrices may not fully capture the complex nature of the neighbour relationship between provinces.

There are other limitations, which can also be attributed to data limitations. The time series of the panel dataset are not long enough to be favourable for studying the long-run growth in Chapter 3. Sectoral disaggregation of private sector employment data is not available for Chapter 4, and therefore the empirical analysis is unable to control for the possibility that the employment impact of infrastructure, as well as of other factors, can vary according to economic sectors. Last but not least, the empirical research emphasises the importance of controlling for unobserved heterogeneity in the intercepts of the models but there may also be heterogeneity in the slope parameters (Lee *et al.*, 1997). More precisely, differences in the fundamentals of provincial economies introduce the possibility of spatial heterogeneity which means that the relationship between growth, or employment, or industrial location choices, and infrastructure (and other factors as well) is not stable over space. However, it is not possible to allow for unrestricted heterogeneity in both the intercepts and the slope coefficients for all the provinces in the dataset used in the empirical research without the availability of longer time series.

6.3.2 Further Research on the Role of Infrastructure

Improvement of the data availability will without any doubt increase the contribution of the empirical research performed in the thesis. A panel dataset with a longer time series will improve the efficiency of the models estimated thanks to an increased number of observations. This in turn allows the estimation using region-based sub-samples of provinces, yielding the empirical results capturing further regional heterogeneities and therefore providing insights for planning the construction of infrastructure at the region level.

With respect to the road data, the dataset containing the composition of roads will allow an estimation of the effects of the interregional and interprovincial transport costs which can be assumed to reduce, and the interregional and interprovincial connectivity which is supposed to be enhanced, along with an improvement in the relevant roads. This research question is important since the interregional and interprovincial transport infrastructure will contribute to making the establishment of more specialised regional economies possible. On the other side, further research on the role of transport infrastructure needs to address the composition of roads and the economic effect of each type of roads installed within a particular province. This research issue is of importance since the availability of transport infrastructure not only for interregional and interprovincial passenger and freight traffics but also for transportation at lower geographical levels, such as district, commune and village, would strengthen the overall transport systems as well as enable the establishment of the multi-level domestic trading system and the achievement of socio-economic outcomes at all geographical levels of the country.

6.3.3 Final Remarks

The role of infrastructure in the economy remains a very broad theme of research. Also, to what extent infrastructure has contributed to the process of economic development in Vietnam is the subject deserved to be explored further in different contexts. The specification of the empirical models for the Vietnamese provinces over the period 2000-2007 is based on the theoretical and empirical relevance, but restricted by the nature and the availability of data. Then, the evaluation of the models should be based on a full consideration of these factors. Finally, it is necessary to emphasise that the empirical findings and their policy implications must be interpreted on an academic basis only.

APPENDIX A: ADDITIONAL RESULTS

Table 1: Tests of Heteroskedasticity and AR(1) for the Employment Model

	(1)	(2)	(3)	(4)	(5)
Heteroskedasticity: χ^2 [*p*-value]	366.17 [0.000]	326.37 [0.000]	328.32 [0.000]	300.89 [0.000]	295.80 [0.000]
First-Order Autocorrelation: *F*-stat [*p*-value]	41.124 [0.000]	56.496 [0.000]	59.211 [0.000]	70.974 [0.000]	67.265 [0.000]

Note: In the heteroskedasticity test, the null hypothesis is there is no heterogeneity across the panels. In the AR(1) test, the null hypothesis is there is no first-order autocorrelation within the panels. The tests are performed for each specification of the employment model presented in Table 4.3.

Table 2: Tests of Heteroskedasticity and AR(1) for the Employment Model with Road Spillover Variable

	(1)	(2)	(3)	(4)	(5)
Heteroskedasticity: χ^2 [*p*-value]	366.54 [0.000]	335.00 [0.000]	338.10 [0.000]	304.88 [0.000]	299.15 [0.000]
First-Order Autocorrelation: *F*-stat [*p*-value]	45.860 [0.000]	72.378 [0.000]	69.378 [0.000]	75.109 [0.000]	71.346 [0.000]

Note: In the heteroskedasticity test, the null hypothesis is there is no heterogeneity across the panels. In the AR(1) test, the null hypothesis is there is no first-order autocorrelation within the panels. The tests are performed for each specification of the employment model with road spillover variable presented in Table 4.4.

Table A.2 (cont.): Tests of Heteroskedasticity and AR(1) for the Employment Model with Road Spillover Variable

	(6)	(7)	(8)	(9)	(10)
Heteroskedasticity: χ^2 [*p*-value]	364.29 [0.000]	340.05 [0.000]	343.44 [0.000]	302.33 [0.000]	300.10 [0.000]
First-Order Autocorrelation: *F*-stat [*p*-value]	43.456 [0.000]	72.549 [0.000]	69.629 [0.000]	72.564 [0.000]	70.547 [0.000]

Table 3: Results of the Poisson Estimation of the Location-Choice Count Regression Model

	Full Sample	Workforce Size based Sub-Samples				Sector based Sub-Samples				
						Manufacturing		Trade		Services
		Micro	Small	Medium	Large	Non-Machinery	Machinery	Wholesale	Retail	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(*ROAD*) $_{i,t-1}$	0.195***	0.200**	0.212***	0.213***	0.129	0.222***	0.326***	0.268***	0.420***	0.273***
	(0.066)	(0.097)	(0.048)	(0.045)	(0.115)	(0.060)	(0.083)	(0.062)	(0.137)	(0.092)
Log(*POPULATION*) $_{i,t-1}$	0.542***	0.582***	0.572***	0.466***	0.175	0.450***	0.791***	0.586***	0.445***	0.671***
	(0.069)	(0.092)	(0.058)	(0.059)	(0.142)	(0.089)	(0.107)	(0.075)	(0.123)	(0.094)
GDPGR $_{i,t-1}$	0.013	0.015	0.004	0.003	-0.016	0.010	0.000	0.017**	0.027*	0.003
	(0.008)	(0.010)	(0.006)	(0.006)	(0.014)	(0.007)	(0.013)	(0.009)	(0.014)	(0.009)
Log(*WAGE*) $_{i,t-1}$	0.580***	0.735***	0.610***	0.057	0.233	0.357**	0.129	0.443**	0.370*	-0.140
	(0.164)	(0.216)	(0.146)	(0.124)	(0.254)	(0.170)	(0.121)	(0.143)	(0.199)	(0.111)
HIGEDU $_{i,t-1}$	0.046***	0.053***	0.033***	0.025***	0.035***	0.038***	0.022***	0.053***	0.051***	0.071***
	(0.004)	(0.006)	(0.003)	(0.003)	(0.009)	(0.004)	(0.004)	(0.005)	(0.008)	(0.007)
NETMIG $_{i,t-1}$	0.007	0.000	0.017***	0.018***	0.014	0.022***	0.020***	0.013***	0.002	0.002
	(0.006)	(0.008)	(0.004)	(0.004)	(0.009)	(0.005)	(0.007)	(0.004)	(0.009)	(0.005)
HERFINDAHL $_{i,t-1}$	-0.858**	-0.725	-0.798***	-1.260***	-1.847*	-1.669***	-2.436***	0.110	-0.077	-1.276**
	(0.376)	(0.582)	(0.264)	(0.383)	(0.971)	(0.511)	(0.857)	(0.552)	(0.723)	(0.535)
LOCALISATION $_{i,t-1}$						0.662***	0.339***	0.391***	0.228***	0.060
						(0.103)	(0.085)	(0.088)	(0.054)	(0.121)
SOECAP $_{i,t-1}$	-0.053	0.185	-0.410***	-0.732***	-1.474***	-0.428***	0.341**	-0.550***	-1.060***	-0.194
	(0.165)	(0.209)	(0.129)	(0.140)	(0.376)	(0.133)	(0.153)	(0.163)	(0.349)	(0.148)
Log(*ESIZE*) $_{i,t-1}$	0.124	-0.071	0.207**	0.853***	1.082***	0.123	1.010***	0.621***	0.300*	0.303***
	(0.099)	(0.118)	(0.103)	(0.087)	(0.157)	(0.095)	(0.144)	(0.110)	(0.173)	(0.104)
IZ $_{i,t-1}$	0.059	0.105	0.048	-0.098*	-0.166	0.146***	-0.112	-0.018	0.171*	0.148**
	(0.058)	(0.082)	(0.052)	(0.058)	(0.112)	(0.057)	(0.092)	(0.092)	(0.092)	(0.074)
Constant	-4.632***	-5.725***	-6.129***	-7.143***	-4.777*	-4.686***	-15.779***	-9.928***	-7.641***	-9.719***
	(1.164)	(1.580)	(1.006)	(1.008)	(2.609)	(1.403)	(1.881)	(1.318)	(1.977)	(1.511)

Log likelihood	-12051.137	-11587.599	-3647.038	-1614.088	-1005.763	-3220.244	-847.041	-3470.448	-4635.194	-1772.524
Wald χ^2 [p-value]	4512.572 [0.000]	2880.784 [0.000]	9015.871 [0.000]	8656.074 [0.000]	1792.457 [0.000]	5205.790 [0.000]	5153.872 [0.000]	10267.654 [0.000]	1098.805 [0.000]	12018.914 [0.000]
Pearson χ^2 [p-value]	21987.67 [0.000]	21373.96 [0.000]	5552.688 [0.000]	1492.886 [0.000]	1040.54 [0.000]	4668.303 [0.000]	848.3938 [0.000]	5508.389 [0.000]	7814.147 [0.000]	2120.911 [0.000]
Observations	366	366	366	366	366	366	366	366	366	366
No. of Provinces	61	61	61	61	61	61	61	61	61	61

Note:

1. The dependent variable is the number of new private sector enterprises.
2. All explanatory variables are lagged for one year. For the estimations of the full sample and workforce size based sub-samples, all explanatory variables are calculated at province level. For the sectoral estimations, the variables $WAGE$, $LOCALISATION$ and $SOECAP$ are calculated at the province-sector level while the rest are calculated at the province level.
3. Year and region dummies are included in all regressions.
4. Robust standard errors are in parentheses.
5. * p<0.100; ** p<0.050; *** p<0.010

Table 4: Results of the Poisson Estimation of the Location-Choice Count Regression Model with w_{cont} Road Spillover Variable

	Full Sample	Workforce Size based Sub-Samples				Sector based Sub-Samples				
		Micro	Small	Medium	Large	Manufacturing		Trade		Services
						Non-Machinery	Machinery	Wholesale	Retail	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(*ROAD*) $_{i,t-1}$	0.194***	0.200**	0.212***	0.213***	0.131	0.213***	0.322***	0.273***	0.419***	0.274***
	(0.066)	(0.098)	(0.048)	(0.045)	(0.114)	(0.061)	(0.082)	(0.064)	(0.135)	(0.094)
Log(w_{cont}*ROAD*) $_{i,t-1}$	-0.013	-0.003	-0.056	0.063	0.180	-0.183*	0.133	-0.181	0.081	-0.154
	(0.093)	(0.127)	(0.089)	(0.101)	(0.176)	(0.109)	(0.158)	(0.121)	(0.176)	(0.144)
Log(*POPULATION*) $_{i,t-1}$	0.543***	0.582***	0.580***	0.457***	0.157	0.457***	0.780***	0.608***	0.436***	0.687***
	(0.071)	(0.094)	(0.060)	(0.062)	(0.144)	(0.088)	(0.108)	(0.077)	(0.127)	(0.096)
GDPGR $_{i,t-1}$	0.013*	0.015	0.004	0.003	-0.015	0.008	0.002	0.016*	0.028**	0.003
	(0.008)	(0.010)	(0.006)	(0.006)	(0.014)	(0.007)	(0.013)	(0.009)	(0.014)	(0.009)
Log(*WAGE*) $_{i,t-1}$	0.578***	0.735***	0.596***	0.071	0.279	0.346**	0.148	0.430***	0.383*	-0.145
	(0.166)	(0.219)	(0.149)	(0.127)	(0.253)	(0.172)	(0.121)	(0.141)	(0.207)	(0.110)
HIGEDU $_{i,t-1}$	0.046***	0.053***	0.033***	0.025***	0.035***	0.038***	0.022***	0.052***	0.051***	0.071***
	(0.004)	(0.006)	(0.003)	(0.003)	(0.009)	(0.004)	(0.004)	(0.005)	(0.008)	(0.007)
NETMIG $_{i,t-1}$	0.007	0.000	0.017***	0.018***	0.014	0.022***	0.020***	0.013***	0.002	0.001
	(0.006)	(0.008)	(0.004)	(0.004)	(0.009)	(0.005)	(0.007)	(0.004)	(0.009)	(0.005)
HERFINDAHL $_{i,t-1}$	-0.871**	-0.729	-0.856***	-1.190***	-1.645	-1.835***	-2.277***	-0.080	0.009	-1.471**
	(0.398)	(0.615)	(0.285)	(0.388)	(1.001)	(0.514)	(0.884)	(0.564)	(0.760)	(0.578)
LOCALISATION $_{i,t-1}$						0.709***	0.339***	0.387***	0.228***	0.044
						(0.106)	(0.084)	(0.089)	(0.053)	(0.123)
SOECAP $_{i,t-1}$	-0.054	0.185	-0.412***	-0.732***	-1.469***	-0.437***	0.369**	-0.545***	-1.035**	-0.185
	(0.164)	(0.209)	(0.129)	(0.139)	(0.375)	(0.133)	(0.158)	(0.160)	(0.347)	(0.148)
Log(*ESIZE*) $_{i,t-1}$	0.125	-0.070	0.210**	0.850***	1.073***	0.126	0.998***	0.627***	0.291*	0.314***
	(0.099)	(0.119)	(0.105)	(0.088)	(0.158)	(0.095)	(0.144)	(0.110)	(0.173)	(0.104)
IZ $_{i,t-1}$	0.059	0.105	0.051	-0.099*	-0.165	0.146***	-0.115	-0.006	0.170*	0.155**
	(0.058)	(0.082)	(0.053)	(0.058)	(0.111)	(0.056)	(0.092)	(0.089)	(0.092)	(0.074)
Constant	-4.552***	-5.703**	-5.811***	-7.508***	-5.938**	-3.388**	-16.588***	-8.938***	-8.104***	-8.842***
	(1.254)	(1.807)	(1.048)	(1.061)	(2.708)	(1.554)	(2.074)	(1.443)	(2.257)	(1.758)

Log likelihood	-12050.659	-11587.579	-3644.203	-1612.956	-1004.214	-3203.006	-846.228	-3457.121	-4632.685	-1768.575
Wald χ^2 [*p*-value]	4521.567 [0.000]	2881.827 [0.000]	9058.816 [0.000]	8741.084 [0.000]	1804.912 [0.000]	5438.794 [0.000]	5078.266 [0.000]	10659.522 [0.000]	1101.071 [0.000]	12209.357 [0.000]
Pearson χ^2 [*p*-value]	21989.65 [0.000]	21373.64 [0.000]	5560.757 [0.000]	1488.09 [0.000]	1033.391 [0.000]	4655.734 [0.000]	853.1334 [0.000]	5460.935 [0.000]	7825.808 [0.000]	2107.161 [0.000]
Observations	366	366	366	366	366	366	366	366	366	366
No. of Provinces	61	61	61	61	61	61	61	61	61	61

Note: See Table A.3

Table 5: Results of the Poisson Estimation of the Location-Choice Count Regression Model with w_{dist} Road Spillover Variable

	Full Sample	Workforce Size based Sub-Samples				Sector based Sub-Samples				
						Manufacturing		Trade		Services
		Micro	Small	Medium	Large	Non-Machinery	Machinery	Wholesale	Retail	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(*ROAD*) $_{i,t-1}$	0.166***	0.173**	0.182***	0.157***	0.067	0.194***	0.234***	0.288***	0.379***	0.265***
	(0.060)	(0.087)	(0.050)	(0.050)	(0.122)	(0.061)	(0.086)	(0.067)	(0.122)	(0.091)
Log(w_{dist}*ROAD*) $_{i,t-1}$	0.142*	0.135	0.154**	0.260***	0.280**	0.153*	0.331**	-0.080	0.174	0.035
	(0.077)	(0.106)	(0.070)	(0.076)	(0.125)	(0.087)	(0.130)	(0.093)	(0.145)	(0.122)
Log(*POPULATION*) $_{i,t-1}$	0.525***	0.566***	0.548***	0.432***	0.146	0.440***	0.753***	0.597***	0.425***	0.666***
	(0.072)	(0.094)	(0.062)	(0.063)	(0.144)	(0.090)	(0.109)	(0.075)	(0.125)	(0.094)
GDPGR $_{i,t-1}$	0.014	0.017	0.006	0.004	-0.013	0.012*	0.006	0.016*	0.030**	0.003
	(0.008)	(0.011)	(0.006)	(0.006)	(0.014)	(0.007)	(0.014)	(0.009)	(0.014)	(0.009)
Log(*WAGE*) $_{i,t-1}$	0.639***	0.786***	0.684***	0.180	0.370	0.384**	0.163	0.442**	0.399**	-0.143
	(0.173)	(0.231)	(0.153)	(0.137)	(0.265)	(0.170)	(0.119)	(0.143)	(0.200)	(0.110)
HIGEDU $_{i,t-1}$	0.046***	0.053***	0.033***	0.025***	0.035***	0.037***	0.023***	0.052***	0.051***	0.071***
	(0.004)	(0.006)	(0.003)	(0.003)	(0.009)	(0.004)	(0.004)	(0.005)	(0.008)	(0.007)
NETMIG $_{i,t-1}$	0.007	0.000	0.016***	0.017***	0.013	0.022***	0.019***	0.013***	0.001	0.001
	(0.006)	(0.008)	(0.004)	(0.004)	(0.009)	(0.005)	(0.006)	(0.004)	(0.009)	(0.005)
HERFINDAHL $_{i,t-1}$	-0.715*	-0.580	-0.667**	-1.000***	-1.554	-1.526***	-1.914**	0.027	0.120	-1.231**
	(0.377)	(0.587)	(0.261)	(0.364)	(0.963)	(0.508)	(0.886)	(0.550)	(0.739)	(0.559)
LOCALISATION $_{i,t-1}$						0.615***	0.366***	0.393***	0.236***	0.065
						(0.113)	(0.077)	(0.089)	(0.053)	(0.123)
SOECAP $_{i,t-1}$	-0.045	0.192	-0.401***	-0.724***	-1.467***	-0.424***	0.405***	-0.543***	-1.023***	-0.197
	(0.164)	(0.209)	(0.127)	(0.135)	(0.376)	(0.133)	(0.155)	(0.161)	(0.346)	(0.150)
Log(*ESIZE*) $_{i,t-1}$	0.104	-0.091	0.190*	0.821***	1.043***	0.113	0.958***	0.628***	0.272	0.298***
	(0.099)	(0.119)	(0.103)	(0.085)	(0.155)	(0.095)	(0.143)	(0.110)	(0.176)	(0.107)
IZ $_{i,t-1}$	0.074	0.119	0.061	-0.067	-0.127	0.165***	-0.078	-0.024	0.194**	0.151**
	(0.057)	(0.081)	(0.050)	(0.053)	(0.108)	(0.055)	(0.090)	(0.090)	(0.094)	(0.075)
Constant	-5.279***	-6.341***	-6.785***	-8.309***	-6.116**	-5.460***	-17.001***	-9.670***	-8.330***	-9.834***
	(1.208)	(1.701)	(0.946)	(0.989)	(2.584)	(1.429)	(1.851)	(1.384)	(2.043)	(1.599)

Log likelihood	-11959.807	-11539.630	-3614.801	-1585.843	-999.989	-3199.910	-838.551	-3466.235	-4617.090	-1772.228
Wald χ^2 [*p*-value]	4566.651 [0.000]	2834.382 [0.000]	10169.544 [0.000]	9110.123 [0.000]	1845.808 [0.000]	5154.506 [0.000]	5242.914 [0.000]	10337.125 [0.000]	1106.357 [0.000]	12027.551 [0.000]
Pearson χ^2 [*p*-value]	21675.7 [0.000]	21209.69 [0.000]	5403.725 [0.000]	1422.792 [0.000]	1027.895 [0.000]	4592.473 [0.000]	819.8504 [0.000]	5493.682 [0.000]	7810.863 [0.000]	2121.95 [0.000]
Observations	366	366	366	366	366	366	366	366	366	366
No. of Provinces	61	61	61	61	61	61	61	61	61	61

Note: See Table A.3

APPENDIX B: THE WAGE MODEL

Similarly to the reduced-form employment equation, the reduced-form equation of private sector wages can be specified as a function of infrastructure and other factors potentially affecting labour demand and labour supply

$$\ln WAGE_{i,t} = \alpha + \beta \ln ROAD_{i,t-1} + \gamma_{LD} LD_{i,t-1} + \delta_{LS} LS_{i,t-1} + \mu_i + \varphi_t + \varepsilon_{i,t} \qquad \text{(B.1)}$$

where $EMP_{i,t}$ is the number of people working in the private sector in province i in year t; $ROAD_{i,t-1}$ is denotes the density of the road network, measured by total road length per 1,000 km^2, in province i in year $t-1$; $LD_{i,t-1}$ is the vector of other factors affecting the private sector labour demand in province i in year $t-1$; $LS_{i,t-1}$ is the vector of other factors affecting the labour supply in province i in year $t-1$; μ_i is the unobserved province-specific factors; φ_t is the unobserved time-specific factors; and $\varepsilon_{i,t}$ is the error terms.

Given the nature of the transition economy of Vietnam in which although economic reforms have led to the scaled-down restructuring of the SOE sector and the explosion of the private sector particularly in terms of the number of non-SOEs, SOEs still maintain a substantial role in the economy, the presence of SOEs in the labour market should be controlled for when estimating the determinants of private sector wages. Then, the wage model is re-written as

$$\ln WAGE_{i,t} = \alpha + \beta \ln ROAD_{i,t-1} + \gamma_{LD} LD_{i,t-1} + \delta_{LS} LS_{i,t-1} + \sigma SOEEMP_{i,t-1} + \mu_i + \varphi_t + \varepsilon_{i,t} \qquad \text{(B.2)}$$

where $SOEEMP_{i,t-1}$ is proportion of people working in SOEs in total employment in province i in year $t-1$.

Similarly to the empirical analysis on employment, the present analysis on wages examines further the wage impact of transport infrastructure by estimating the wage model which incorporates both own-province and neighbouring provinces' transport infrastructure

$$\ln WAGE_{i,t} = \alpha + \beta \ln ROAD_{i,t-1} + \gamma_{LD} LD_{i,t-1} + \delta_{LS} LS_{i,t-1} + \sigma SOEEMP_{i,t-1} + \omega \ln w_{i,j} ROAD_{i,t-1} + \mu_i + \varphi_t + \varepsilon_{i,t} \quad \text{(B.3)}$$

where $w_{i,j}ROAD_{i,t-1}$ is the average road density of neighbouring provinces j of province i in year $t-1$.

Results of the wage model are presented in Table B.1. Wald χ^2 is reported as a diagnostic tool for the validity of the FGLS regression, which is statistically significant at the 1 per cent level in all cases. Also, the problems of heteroskedasticity and panel-specific first-order autocorrelation are controlled for in all regressions.

Table 6: FGLS Results of the Wage Model

	(1)	(2)	(3)	(4)	(5)
Log(*ROAD*) $_{i,t-1}$	0.188***	0.224***	0.223***	0.220***	0.166**
	(0.056)	(0.061)	(0.061)	(0.061)	(0.065)
Log(*OUTPUT*) $_{i,t-1}$	0.125***				
	(0.017)				
Log(*CAPITAL*) $_{i,t-1}$		0.063***	0.072***		
		(0.014)	(0.015)		
ΔLog(*SOLOW*) $_{i,t-1}$			0.028**	0.032**	0.031**
			(0.013)	(0.013)	(0.013)
Log(*CAPITALPE*) $_{i,t-1}$				0.086***	0.080***
				(0.014)	(0.015)
Log(*ENTERPRISES*) $_{i,t-1}$				0.016	0.030
				(0.025)	(0.024)
NETMIG $_{i,t-1}$	0.005***	0.003***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log(*URBAN*) $_{i,t-1}$	-0.009	-0.032	-0.037	-0.030	-0.126
	(0.064)	(0.079)	(0.079)	(0.079)	(0.082)
ΔLog(*SECGRA*) $_{i,t-1}$	-0.033**	-0.025	-0.023	-0.026	-0.027
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
HIGEDU $_{i,t-1}$	0.000	-0.001	-0.001	0.000	0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Log(*POPULATION*) $_{i,t-1}$					-0.674***
					(0.184)
SOEEMP $_{i,t-1}$	0.023	0.001	0.018	-0.045	-0.010
	(0.071)	(0.081)	(0.079)	(0.085)	(0.089)
Constant	-1.692***	-0.747	-0.903	-0.503	9.653***
	(0.597)	(0.579)	(0.561)	(0.595)	(2.851)
Wald χ^2 [*p*-value]	82.03 [0.000]	52.77 [0.000]	60.13 [0.000]	85.59 [0.000]	78.54 [0.000]
Observations	366	366	366	366	366
No. of Provinces	61	61	61	61	61

Note:

1. The dependent variable is the average real wages per private sector employee, Log(*WAGE*).
2. All explanatory variables are lagged for one year.
3. Province and year dummies are included in all regressions.
4. Standard errors are in parentheses.
5. * p<0.100; ** p<0.050; *** p<0.010
6. The problems of heteroskedasticity and autocorrelation are controlled for in all regressions.

The coefficient on the variable $ROAD$ is positive and statistically significant in all regressions. Throughout various specifications of the wage model, transport infrastructure shows the largest positive effect on the wages paid in the private sector. The magnitude of the coefficient is at around 0.2, which suggests that a 1 per cent increase in the measure of transport infrastructure leads to a 0.2 per cent increase in private sector wages.

The real average wages per employee paid in the private sector is also driven by an expansion of the production process. This is indicated by the coefficient on the variable $OUTPUT$ which is positive and statistically significant at the 1 per cent level. The magnitude of the coefficient is 0.125, which suggests that a 1 per cent increase in private sector output causes an increase of private sector wages by 0.125 per cent.

As stated previously, when enterprises increase their physical capital stock, two competing effects might occur. On the one side, enterprises might substitute physical capital for labour which would lead to a decrease in the labour demand. Given that there is no decline in the labour supply, this would in turn lead to a decrease in the wage rate. On the other side, if an increase in physical capital stock accumulation leads to an increase in productivity, and given that the amount of labour input required to produce the amount of output demanded remains unchanged, that the increased capital accumulation might have the consequence of a higher wage rate. The issue becomes more complex since the composition of skilled and unskilled workers is unknown in the present study. In the employment analysis presented in Chapter 4, the physical capital variables, i.e. the stock, $CAPITAL$, and the stock per enterprise, $CAPITALPE$, are statistically insignificant in all regressions. As the dependent variable is the number of workers employed in the private sector, there is no skill premium and this could provide another explanation on why these variables do not act as important determinants of employment. However, the effect of physical capital stock accumulation appears to be more apparent in the wage analysis. The magnitude of the coefficient on the physical capital stock per enterprise variable is even larger than those of productivity growth and the number of enterprises. Accordingly, a 1 per cent increase in physical capital stock per private sector enterprise raises private sector wages by around 0.08 per cent.

The coefficient on the productivity growth variable, $\Delta SOLOW$, is positive and statistically significant at the 5 per cent level in all regressions where it is added, suggesting that the growth of productivity contributes to an increase in wages paid to private sector employees. The magnitude of the coefficient is at around 0.03, which indicates that a 1 per cent increase in the measure of private sector productivity growth is associated with an increase by around 0.03 per cent in private sector wages.

For a given labour supply, an increase in the number of enterprises located in a particular province is expected to result in an increase in the labour demand and therefore an increase in wages. However, in the present study, the coefficient on the number of private sector enterprises, $ENTERPRISES$, is statistically insignificant, which might imply that the competition for labour was not be severe enough to lead to enterprises paying more for their employees. Alternatively, the increased labour demand induced by the increased number of enterprises but also accompanied by a simultaneous increase in labour supply would not cause wages to increase. The stimulations increases in labour demand and supply might happen over the sample period which witnessed the explosion of the number of private sector enterprises as well as an annually increased labour participation rate.

Labour mobility would provide another explanation for the finding that an increase in the number of private sector enterprises is not a key determinant of an increase in private sector wages. The effect of the net migration rate appears to be more apparent in the wage analysis than that in the employment analysis. As argued previously, an increase in the supply of labour in-migrating from elsewhere does not necessary encourage enterprises to expand their workforce. However, immigrants might contribute to bidding up the costs of living in some particular areas, which is in part explained by congestion in using public goods and other basic amenities as well as congestion in sharing such immobile factors as land and housing, and the wage rate is consequently required to be raised up so as to cover the increased living costs.

The variables aimed at capturing the supply of qualified labour, i.e. the change of the number of high-school graduates, $\Delta SECGRA$, and the endowment of higher-education infrastructure, $HIGEDU$, are positively associated with private sector employment, but do not affect private sector wages. One explanation could be that an increased supply of qualified labour would make the competition for labour in general and skilled labour in particular less severe, and this in turn would reduce the pressure on private sector enterprises to raise the real average wages paid to their employees.

The coefficient on the variable $URBAN$ is statistically insignificant in all regressions, suggesting that urbanisation does not appear to be important for wages in the private sector while, as presented in Chapter 4, it is an important determinant of private sector employment. One explanation could be that a more urbanised province would provide fewer alternative options available to job seekers, while providing more alternative options available to employers, and therefore a wage premium does not have to be paid to private sector employees.

Population appears to be the key factor explaining a decrease in the private sector wages while it does not have any impact on employment in the private sector. The results for the wage model show that the coefficient on the variable $POPULATION$ is negative and statistically significant at the 1 per cent level. As noted previously, interpretation of the result for this variable requires an assumption that the age structure of the population is homogenous across provinces. Then, an increase in the total number of inhabitants in a particular province would lead to an increase in the labour supply and consequently, for a given labour demand, a decrease in wages paid to private sector employees. The magnitude of the coefficient, which is around -0.6, suggests that a 1 per cent increase in the number of inhabitants leads to a decrease of around 0.6 per cent in private sector wages.

One might expect that an increase in the proportion of SOE employment in total employment would intensify the competition for labour and, consequently, the private sector has to raise their average wages to compete with SOEs for labour. However, in the present study, the variable $SOEEMP$ does not have a statistically significant impact on private sector wages. This could be due to the fact that there is no information on the skill premium of the two sectors. Given the large pool of labour, in particular unskilled labour, of Vietnam over the sample period, such the competition might not be severe enough to raise wages paid in the private sector. In practice, the competition between the two sectors for labour might be driven by not only wage differences but also by, for example, differences in regard to working conditions which, however, cannot be controlled for in the present study. Another explanation could be that the competition between SOEs and non-SOEs not only take places in the formal labour market, but also in other markets such as the capital market and the markets for goods produced by both sectors. Such the competition affects negatively private sector production and hence employment, but not necessarily leads to a decrease in private sector wages.

Results reported in Table B.2 show a more substantial role of neighbouring roads compared with that of roads installed in the home province in explaining private sector wages. The coefficient on the road spillover variable is positive and statistically significant at the 1 per cent level in all regressions, indicating the positive relationship between wages paid in a particular province and transport infrastructure installed in its neighbouring provinces. The magnitude of the coefficient is at around 0.3, which suggests that a 1 per cent increase in the measure of transport infrastructure spillovers leads to an increase in private sector wages by around 0.3 per cent. One explanation for the positive spillover effect is that an increase in infrastructure endowment should increase the attractiveness of the neighbouring provinces to residential and industrial location choices as well as the competitiveness of enterprises located in those provinces and therefore enterprises located in a particular province must increase the wage rate in order to compete with the neighbouring enterprises for labour.

Table 7: FGLS Results of the Wage Model with Road Spillover Variable

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(*ROAD*) $_{i,t-1}$	0.149**	0.207***	0.207***	0.192***	0.145**	0.156***	0.214***	0.213***	0.201***	0.146**
	(0.058)	(0.061)	(0.061)	(0.062)	(0.065)	(0.058)	(0.061)	(0.061)	(0.062)	(0.065)
Log(w_{cont}*ROAD*) $_{i,t-1}$	0.280***	0.316***	0.280***	0.307***	0.305***					
	(0.084)	(0.072)	(0.076)	(0.082)	(0.082)					
Log(w_{dist}*ROAD*) $_{i,t-1}$						0.259***	0.280***	0.246***	0.281***	0.282***
						(0.084)	(0.067)	(0.072)	(0.080)	(0.078)
Log(*OUTPUT*) $_{i,t-1}$	0.126***					0.126***				
	(0.019)					(0.018)				
Log(*CAPITAL*) $_{i,t-1}$		0.060***	0.066***				0.057***	0.064***		
		(0.014)	(0.014)				(0.014)	(0.014)		
ΔLog(*SOLOW*) $_{i,t-1}$			0.021	0.024*	0.025*			0.021*	0.025*	0.026**
			(0.013)	(0.013)	(0.013)			(0.013)	(0.013)	(0.013)
Log(*CAPITALPE*) $_{i,t-1}$				0.077***	0.074***				0.078***	0.073***
				(0.014)	(0.015)				(0.014)	(0.015)
Log(*ENTERPRISES*) $_{i,t-1}$				0.014	0.026				0.011	0.023
				(0.024)	(0.023)				(0.024)	(0.023)
NETMIG $_{i,t-1}$	0.005***	0.004***	0.003***	0.003***	0.004***	0.005***	0.004***	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log(*URBAN*) $_{i,t-1}$	-0.010	-0.054	-0.053	-0.063	-0.146*	-0.011	-0.050	-0.050	-0.058	-0.146*
	(0.076)	(0.083)	(0.083)	(0.084)	(0.086)	(0.072)	(0.080)	(0.081)	(0.081)	(0.085)
ΔLog(*SECGRA*) $_{i,t-1}$	-0.034**	-0.024	-0.022	-0.024	-0.025	-0.033*	-0.024	-0.022	-0.025	-0.026
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
HIGEDU $_{i,t-1}$	0.000	-0.001	-0.001	0.000	0.003	-0.000	-0.002	-0.001	-0.000	0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Log(*POPULATION*) $_{i,t-1}$					-0.660***					-0.679***
					(0.183)					(0.186)
SOEEMP $_{i,t-1}$	0.027	0.008	0.015	-0.055	-0.036	0.020	-0.006	0.005	-0.057	-0.044
	(0.073)	(0.079)	(0.078)	(0.085)	(0.090)	(0.072)	(0.080)	(0.079)	(0.085)	(0.090)
Constant	-3.459***	-2.900***	-2.757***	-2.466***	7.467***	-3.306***	-2.575***	-2.466***	-2.285***	7.999***
	(0.781)	(0.714)	(0.715)	(0.730)	(2.830)	(0.782)	(0.702)	(0.707)	(0.733)	(2.870)

Wald χ^2 [*p*-value]	92.12 [0.000]	82.41 [0.000]	83.57 [0.000]	117.69 [0.000]	90.11 [0.000]	93.55 [0.000]	78.61 [0.000]	78.02 [0.000]	105.34 [0.000]	87.23 [0.000]
Observations	366	366	366	366	366	366	366	366	366	366
No. of Provinces	61	61	61	61	61	61	61	61	61	61

Note: See Table B.1

APPENDIX C: THE DISCRETE LOCATION-CHOICE MODEL

The expected profit of the representative enterprise n can be specified as

$$\pi_{n,i,t} = \alpha + \beta' X_{n,i,t-1} + \gamma_r \sum_{r=1}^{5} R_{n,r} + \varepsilon_{n,i,t} \tag{C.1}$$

where $\pi_{n,i,t}$ is the expected profit of the representative enterprise n when establishing in province i in year t; $X_{n,i,t-1}$ is the vector of explanatory variables capturing the specific characteristics of province i in year $t-1$; $R_{n,r}$ is the vector of region dummies; and $\varepsilon_{n,i,t}$ is the error terms.

The vector X is similar to the vector of local attributes examined in Chapter 5. Then, the expected profit equation can also be written as

$$\begin{aligned}\pi_{n,i,t} = \alpha + \beta \ln ROAD_{n,i,t-1} + \delta_D D_{n,i,t-1} + \mu_L L_{n,i,t-1} + \psi_A A_{n,i,t-1} + \omega_G G_{n,i,t-1} \\ + \gamma_r \sum_{r=1}^{5} R_{n,r} + \varepsilon_{n,i,t}\end{aligned} \tag{C.2}$$

where $ROAD_{n,i,t-1}$ denotes the density of the road network, measured by total road length per 1,000 km^2, in province i in year $t-1$; $D_{n,i,t-1}$ is the vector of variables capturing local demand in province i in year $t-1$; $L_{n,i,t-1}$ is the vector of variables capturing wages and labour supply in province i in year $t-1$; $A_{n,i,t-1}$ is the vector of variables capturing agglomeration economies in province i in year $t-1$; and, $G_{n,i,t-1}$ is the vector of variables capturing the implementation of public policies in province i in year $t-1$.

Province i is chosen if and only if

$$\pi_{n,i,t} > \pi_{n,I,t}, \forall i \neq I \tag{C.3}$$

where I indexes all 61 location-choice alternatives, i.e. 61 provinces, available to the representative firm n.

The probability of the representative enterprise choosing a particular province i out of 61 potential provinces can be expressed as

$$\text{Prob}(n,i) = \frac{\exp\left(\beta' X_{n,i,t-1} + \gamma' \sum_{r=1}^{5} R_{n,r}\right)}{\sum_{s=1}^{61} \exp\left(\beta' X_{n,s,t-1} + \gamma' \sum_{r=1}^{5} R_{n,r}\right)} \tag{C.4}$$

It is worth noting that, in the Conditional Logit model, the specific characteristics of each enterprise cannot be included as covariates. Since the characteristics of each chooser do not vary between location-choice alternatives, i.e. provinces, they cannot determine which province is chosen. Some studies add interaction terms between a variable representing location-specific characteristics and a variable representing firm-specific characteristics. In another way, some studies control for the fact that the specific characteristics of firms could influence their location choices by estimating separately sub-samples containing firms that have similar workforce-size or industry.

Results of the Conditional Logit estimation are presented in Table C.1. In all cases, the likelihood ratio index ρ^2, or the Pseudo R^2, which is the goodness-of-fit indicator of the Conditional Logit model, is satisfactory and the likelihood ratio χ^2 test rejects the null hypothesis that all the coefficients are zero. According to Louviere et al. (2000, p.55), ρ^2 that ranges from 0.2 to 0.4 is considered “extremely good model fits”.

Table 8: Results of the Conditional Logit Estimation of the Discrete Location-Choice Model

	Manufacturing		Trade		Services
	Non-Machinery	Machinery	Wholesale	Retail	
	(1)	(2)	(3)	(4)	(5)
Log(*ROAD*) $_{i,t-1}$	0.221***	0.338***	0.267***	0.420***	0.272***
	(0.015)	(0.051)	(0.015)	(0.016)	(0.028)
Log(*POPULATION*) $_{i,t-1}$	0.449***	0.789***	0.586***	0.444***	0.670***
	(0.020)	(0.070)	(0.020)	(0.022)	(0.033)
GDPGR $_{i,t-1}$	0.010***	-0.001	0.017***	0.027	0.003
	(0.002)	(0.008)	(0.002)	(0.002)	(0.003)
Log(*WAGE*) $_{i,t-1}$	0.357***	0.140*	0.442***	0.370***	-0.140***
	(0.032)	(0.072)	(0.027)	(0.034)	(0.041)
HIGEDU $_{i,t-1}$	0.037***	0.021***	0.052***	0.051***	0.070***
	(0.000)	(0.002)	(0.000)	(0.000)	(0.002)
NETMIG $_{i,t-1}$	0.022***	0.019***	0.013***	0.001	0.001
	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
HERFINDAHL $_{i,t-1}$	-1.668***	-1.999***	0.109	-0.076	-1.275***
	(0.114)	(0.513)	(0.124)	(0.123)	(0.241)
LOCALISATION $_{i,t-1}$	0.662***	0.332***	0.390***	0.227***	0.060
	(0.024)	(0.041)	(0.014)	(0.010)	(0.039)
SOECAP $_{i,t-1}$	-0.427***	0.306***	-0.549***	-1.059***	-0.193***
	(0.034)	(0.090)	(0.038)	(0.048)	(0.052)
Log(*ESIZE*) $_{i,t-1}$	0.122***	0.978***	0.620***	0.299***	0.303***
	(0.020)	(0.076)	(0.025)	(0.029)	(0.038)
IZ $_{i,t-1}$	0.146***	-0.104*	-0.017	0.170***	0.148***
	(0.014)	(0.055)	(0.019)	(0.019)	(0.029)
Pseudo R^2	0.2332	0.3395	0.3957	0.2349	0.4820
Log likelihood	-116276.88	-9906.4984	-102582.11	-76160.672	-41727.476
LR χ^2 [*p*-value]	70738.30 [0.000]	10183.92 [0.000]	134344.64 [0.000]	46776.50 [0.000]	77650.20 [0.000]
Observations	2250229	208656	2518934	1477176	1195295
No. of Provinces	61	56	61	61	61

Note:

1. The dependent variable is a binary variable taking 1 if the observed enterprise chooses to locate in the province i and 0 otherwise.
2. All explanatory variables are lagged for one year. For the sectoral estimations, the variables $WAGE$, $LOCALISATION$ and $SOECAP$ are calculated at the province-sector level while the rest are calculated at the province level.
3. Region dummies are included in all regressions.
4. Standard errors are in parentheses.
5. * $p<0.100$; ** $p<0.050$; *** $p<0.010$

However, results for the Hausman-McFadden test presented in Table C.2 suggest the violation of the IIA assumption at the significant level of 1 per cent. On the one hand, that violation indicates that the model might not be well specified and, hence, might generate improper estimation. On the other hand, it is worth recalling that the Conditional Logit model is suggested to be adopted only in the cases where the alternatives, i.e. provinces, "can plausibly be assumed to be distinct and weighed independently in the eyes of each decision maker" (McFadden, 1974). Furthermore, as noted in Chapter 5, it is acknowledged that the IIA assumption is often violated in the location choice analysis (e.g., Cheng, 2006; Nefussi and Schwellnus, 2010).

Table 9: Results of the Hausman-McFadden Test of IIA Assumption

Omitted Province	Manufacturing		Trade		Services
	Non-Machinery	Machinery	Wholesale	Retail	
	(1)	(2)	(3)	(4)	(5)
1	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.0000	(-)	0.0000	(-)	(-)
3	0.0000	0.0525	0.0262	0.2177	(-)
4	0.0000	0.0274	0.0026	0.0008	0.6997
5	0.4917	0.8464	0.0000	0.0000	0.0003
6	0.9995	0.0000	(-)	0.0837	0.0366
7	0.0000	0.9034	0.0000	0.0000	0.0001
8	0.0000	0.9967	0.0000	(-)	0.8688
9	(-)	0.3392	0.0000	(-)	(-)
10	0.2841	0.7109	0.0206	0.0207	0.6162
11	0.0000	0.2971	(-3.51)	0.9935	0.9084
12	0.0000	0.9584	0.0164	(-)	(-)
13	0.4174	0.9911	0.0000	(-)	0.9981
14	0.9972	0.9785	0.0000	1.0000	0.9998
15	(-)	NA	0.9990	0.8556	1.0000
16	0.0000	(-)	0.0008	0.0000	0.0004
17	0.0202	1.0000	0.0743	0.0281	0.9791
18	0.1337	0.9946	0.0149	0.9997	0.9785
19	0.0000	0.9836	(-)	0.9965	0.8207
20	0.1206	0.9098	0.0022	1.0000	1.0000
21	0.9998	1.0000	0.9668	0.9946	1.0000
22	0.2398	0.9962	0.0000	0.0126	0.2955
23	1.0000	1.0000	0.0177	0.6258	0.5121
24	0.0005	NA	0.0731	0.0000	0.2516

25	0.9913	(-)	0.9885	0.5299	0.9578
26	0.0003	0.1055	0.0000	0.0000	0.0055
27	0.9974	1.0000	(-)	0.9188	0.0000
28	0.0277	0.9808	(-)	0.0000	(-)
29	(-)	0.8723	0.0000	0.0000	0.1996
30	0.9987	0.9995	0.0000	0.0000	1.0000
31	0.9649	0.9993	0.4014	0.0154	(-)
32	0.0000	0.9897	0.0000	0.0000	0.0000
33	(-)	1.0000	0.0000	0.0000	0.0000
34	0.0000	1.0000	0.0000	0.0000	(-)
35	0.1236	0.9928	0.0000	(-)	0.0424
36	0.3252	1.0000	0.0000	0.8420	0.9955
37	(-)	0.9999	0.0000	0.2456	0.9075
38	0.0000	1.0000	0.0000	0.0000	0.1536
39	0.0000	1.0000	0.7436	0.0000	0.8963
40	0.7737	NA	0.8786	(-)	(-)
41	0.9595	0.9999	0.0023	(-)	(-)
42	0.5676	1.0000	0.0000	0.0784	0.9975
43	0.9603	1.0000	1.0000	0.0000	0.0189
44	0.9344	0.2161	0.0046	0.0000	0.1374
45	0.0000	0.1230	0.7341	0.0001	(-)
46	0.0000	0.1855	0.0113	0.0000	0.6320
47	0.0000	1.0000	0.0025	0.0000	0.0000
48	0.0000	0.9427	0.0000	0.0000	(-)
49	0.0000	0.0142	0.0000	0.0000	(-)
50	0.0000	0.9996	0.0000	(-)	(-)
51	0.0000	1.0000	0.6142	(-)	0.0000
52	(-)	0.8522	(-)	0.0000	(-)
53	0.2331	(-)	0.6798	0.0000	0.5249
54	0.0000	0.7849	0.0001	0.0000	0.2849
55	0.0294	NA	0.0300	0.1826	0.0003
56	0.0000	0.9989	1.0000	0.0288	0.0000
57	0.0012	0.3332	0.3935	0.0000	0.0342
58	0.0000	(-)	(-)	0.0000	(-)
59	0.0000	NA	0.8452	0.0000	0.0000
60	0.9396	1.0000	(-)	0.0000	0.9259
61	(-)	1.0000	0.0000	0.0000	0.9998

Note:

1. The test is carried out for the regressions reported in Table C.1.
2. The null hypothesis is that the IIA assumption holds.
3. *p*-values are reported.
4. (-) denotes a negative chi-square.
5. NA denotes the province that did not have any new enterprises over the sample period and, hence, was not included in the estimation.

The Random Utility Maximisation Nested Logit model (henceforth, the Nested Logit model) is adopted as an alternative to the Conditional Logit model. In particular, the Nested Logit model considers the location choice to be a hierarchical process. To construct the region-based nest structure, the present study follows the geographical classification of 61 provinces into 6 regions, which can be seen in various publications of the General Statistics Office of Vietnam. Those 6 regions include the Red River Delta, RRD, the Northern Mountain, NM, the Central Coast, CC, the Central Highlands, CH, the Southeast, SE, and the Mekong River Delta, MRD. However, results of the Nested Logit estimation, reported in Table C.3, suggest that the region-based hierarchical location choice structure is rejected in all cases. Specifically, the results show that although the inclusive value coefficients are statistically significant at the 1 per cent level, they are not in the range between zero and one. The statistically significant inclusive value coefficients indicate that the two levels of the Nested Logit model, i.e. regions and provinces, are related. However, the inclusive value coefficients that are above 1 imply that the model is inconsistent with RUM. As a result, the region-based hierarchical location choice structure cannot be empirically confirmed.[44]

An alternative nest structure is constructed basing on the average share of the province in the national real GDP over the period 2000-2007. In particular, provinces are categorised into 3 GDP-based zones. The first zone, $ZONE1$, contains Ha Noi, Ho Chi Minh City and Ba Ria - Vung Tau, which accounted for the average share of 7 per cent per year, 18 per cent per year, and 8 per cent per year, respectively, in the national GDP. The second zone, $ZONE2$, contains 17 provinces, and each province had the average contribution to the national GDP ranging from 2 per cent per year to 4 per cent per year. The third zone, $ZONE3$, contains 41 provinces, and each province contributed to the average of around 1 per cent per year or less than 0.5 per

[44] According to McFadden (1978), the inclusive value coefficient represents the degree of independence between alternatives, i.e. provinces, within the same nest, i.e. region, and its value needs to be in the range of between zero and one to be consistent with RUM. The statistically significant coefficients of the inclusive values would indicate that the hierarchical structure of location choice is highly relevant (Hansen, 1987).

cent per year to the national GDP. Results presented in Table C.4 indicate that the GDP-based hierarchical location choice structure can be empirically confirmed only for the machinery and retail sectors. This is suggested by the inclusive value coefficients which are in the expected range and statistically significant at the 1 per cent level. However, this GDP-based nest structure can be criticised as arbitrary.

Generally, neither the Conditional Logit model nor, in most cases, the Nested Logit model produce results that are statistically sound. As Greene (2008b), among others, suggests, alternative methods that do not require the strict assumption of IIA should be adopted in the location choice analysis. The present study uses the count regression models as an alternative to the Conditional Logit and the Nested Logit models, and the empirical findings have been presented and discussed in Chapter 5.

Table 10: Results of the Nested Logit Estimation of the Discrete Location-Choice Model using Region-based Nest Structure

	Manufacturing		Trade		Services
	Non-Machinery	Machinery	Wholesale	Retail	
	(1)	(2)	(3)	(4)	(5)
Log($ROAD$) $_{i,t-1}$	0.304***	0.652***	0.775***	3.919***	0.484***
	(0.020)	(0.087)	(0.066)	(0.413)	(0.067)
Log($POPULATION$) $_{i,t-1}$	0.556***	1.537***	1.570***	5.846***	1.308***
	(0.024)	(0.142)	(0.094)	(0.502)	(0.107)
$GDPGR$ $_{i,t-1}$	0.012***	0.016	0.025***	0.074***	0.010**
	(0.002)	(0.011)	(0.004)	(0.013)	(0.004)
Log($WAGE$) $_{i,t-1}$	0.404***	0.319***	0.974***	0.594***	-0.074
	(0.040)	(0.097)	(0.089)	(0.145)	(0.058)
$HIGEDU$ $_{i,t-1}$	0.041***	0.017***	0.098***	0.211***	0.080***
	(0.001)	(0.003)	(0.006)	(0.012)	(0.005)
$NETMIG$ $_{i,t-1}$	0.023***	0.030***	0.019***	-0.000	0.006***
	(0.001)	(0.004)	(0.001)	(0.003)	(0.002)
$HERFINDAHL$ $_{i,t-1}$	-1.819***	-1.279*	-0.630*	-6.224***	-1.428***
	(0.162)	(0.677)	(0.358)	(1.111)	(0.392)
$LOCALISATION$ $_{i,t-1}$	0.766***	0.407***	0.855***	1.129***	-0.101**
	(0.033)	(0.065)	(0.056)	(0.082)	(0.044)
$SOECAP$ $_{i,t-1}$	-0.536***	0.412***	-1.307***	-3.491***	-0.202**
	(0.040)	(0.140)	(0.092)	(0.239)	(0.085)
Log($ESIZE$) $_{i,t-1}$	0.113***	1.092***	1.418***	1.510***	0.315***
	(0.023)	(0.128)	(0.096)	(0.217)	(0.057)
IZ $_{i,t-1}$	0.181***	0.172**	0.031***	-0.467***	0.289***
	(0.016)	(0.070)	(0.044)	(0.161)	(0.046)
λ_{RRD}	1.093***	1.282***	1.994***	6.384***	1.191***
	(0.032)	(0.103)	(0.120)	(0.483)	(0.081)
λ_{NM}	1.101***	1.532***	3.011***	10.754***	1.503***
	(0.050)	(0.217)	(0.242)	(0.911)	(0.191)
λ_{CC}	1.220***	1.516***	3.034***	9.787***	1.778***
	(0.045)	(0.176)	(0.215)	(0.778)	(0.163)
λ_{CH}	1.343***	1.627***	4.959***	20.190***	2.675***
	(0.100)	(0.544)	(0.440)	(1.795)	(0.373)
λ_{SE}	1.156***	1.334***	2.151***	7.877***	1.333***
	(0.030)	(0.116)	(0.135)	(0.637)	(0.096)
λ_{MRD}	1.205***	1.601***	3.055***	9.295***	1.592***
	(0.046)	(0.221)	(0.228)	(0.753)	(0.176)
Log likelihood	-116309.54	-9926.3577	-102538.94	-76264.872	-41795.099
LR χ^2 [p-value]	70672.98 [0.000]	10144.21 [0.000]	134431 [0.000]	46568.1 [0.000]	77514.95 [0.000]
Observations	2250229	208656	2518934	1477176	1195295
No. of Provinces	61	56	61	61	61

Note:

1. The dependent variable is the binary variable taking 1 if the enterprise chooses to locate in the province i and 0 otherwise.
2. All explanatory variables are lagged for one year. For the sectoral estimations, the variables $WAGE$, $LOCALISATION$ and $SOECAP$ are calculated at the province-sector level while the rest are calculated at the province level.
3. λ denotes the inclusive value.
4. Standard errors are in parentheses.
5. * p<0.100; ** p<0.050; *** p<0.010

Table 11: Results of the Nested Logit Estimation of the Discrete Location-Choice Model using GDP-based Nest Structure

	Manufacturing		Trade		Services
	Non-Machinery	Machinery	Wholesale	Retail	
	(1)	(2)	(3)	(4)	(5)
Log($ROAD$) $_{i,t-1}$	0.313***	0.425***	0.462***	0.232***	0.135**
	(0.048)	(0.063)	(0.048)	(0.041)	(0.061)
Log($POPULATION$) $_{i,t-1}$	0.511***	0.817***	1.717***	0.513***	1.960***
	(0.074)	(0.091)	(0.129)	(0.101)	(0.286)
$GDPGR$ $_{i,t-1}$	0.018***	0.002	0.065***	0.019***	0.050***
	(0.006)	(0.006)	(0.005)	(0.007)	(0.010)
Log($WAGE$) $_{i,t-1}$	0.173***	0.208***	1.047***	0.482***	0.229***
	(0.044)	(0.054)	(0.088)	(0.074)	(0.077)
$HIGEDU$ $_{i,t-1}$	0.066***	0.019***	0.101***	0.041***	0.126***
	(0.009)	(0.002)	(0.006)	(0.007)	(0.012)
$NETMIG$ $_{i,t-1}$	0.027***	0.019***	0.029***	0.014***	0.013***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.003)
$HERFINDAHL$ $_{i,t-1}$	-1.744***	-0.652*	-1.304***	-0.362***	-5.011***
	(0.274)	(0.350)	(0.312)	(0.124)	(0.654)
$LOCALISATION$ $_{i,t-1}$	0.985***	0.185***	0.763***	0.233***	-0.138*
	(0.159)	(0.043)	(0.054)	(0.041)	(0.075)
$SOECAP$ $_{i,t-1}$	-0.562***	0.244***	-0.849***	-0.527***	-0.398**
	(0.078)	(0.078)	(0.078)	(0.090)	(0.168)
Log($ESIZE$) $_{i,t-1}$	0.028	0.583***	1.286***	-0.183***	0.596***
	(0.024)	(0.079)	(0.102)	(0.040)	(0.131)
IZ $_{i,t-1}$	0.429***	0.259***	0.013***	0.458***	0.429***
	(0.062)	(0.056)	(0.049)	(0.091)	(0.093)
λ_{ZONE1}	1.075***	0.571***	2.399***	0.543***	1.987***
	(0.153)	(0.060)	(0.158)	(0.093)	(0.235)
λ_{ZONE2}	1.479***	0.765***	2.855***	0.863***	2.641***
	(0.218)	(0.103)	(0.223)	(0.180)	(0.370)
λ_{ZONE3}	1.319***	0.721***	2.613***	0.837***	2.439***
	(0.197)	(0.102)	(0.199)	(0.164)	(0.339)
Log likelihood	-116111.02	-9915.5416	-102783.62	-77122.41	-42016.06
LR χ^2 [p-value]	71070.01 [0.000]	10165.84 [0.000]	133941.6 [0.000]	44853.02 [0.000]	77073.03 [0.000]
Observations	2250229	208656	2518934	1477176	1195295
No. of Provinces	61	56	61	61	61

Note: See Table C.3

BIBLIOGRAPHY

Abreu, M., De Groot, H., & Florax, R. (2005). Space and Growth: A Survey of Empirical Evidence and Methods. *Region et Development, 21*(1), 13-44.

Alañón-Pardo, A., & Arauzo-Carod, J. (2011). Agglomeration, Accessibility and Industrial Location: Evidence from Spanish Municipalities. Documentos de trabajo de la Facultad de Ciencias Económicas y Empresariales, Universidad Complutense de Madrid.

Allison, P., & Waterman, R. (2002). Fixed-effects Negative Binomial Regression Models. *Socialogical Methodology, 32*(1), 247-265.

Alonso-Borrego, C., & Arellano, M. (1999). Symmetrically Normalised Instrumental-Variable Estimation Using Panel Data. *Journal of Business and Economic Statistics, 17*(1), 36-49.

Álvarez, A., Arias, C., & Orea, L. (2006). Econometric Testing of Spatial Productivity Spillovers from Public Capital. *Revista de Economía Pública, 178*(3), 9-21.

Annez, P., & Buckley, R. (2008). Urbanisation and Growth: Setting the Context. In P. Annez, R. Buckley, & M. Spence (Eds.), *Urbanisation and Growth* (pp. 1-45). Washington, DC: World Bank.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models.* Dordrecht: Kluwer Academic Publishers.

Anselin, L. (2001). Spatial Econometrics. In B. Baltagi (Ed.), *A Companion to Theoretical Econometrics* (pp. 310-330). Oxford: Basil Blackwell.

Anselin, L. (2003). Spatial Externalities, Spatial Multipliers and Spatial Econometrics. *International Regional Science Review, 26*, 153-166.

Anselin, L., & Bera, A. (1998). Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. In A. Ullag, & D. Giles (Eds.), *Handbook of Applied Economic Statistics* (pp. 237-289). Berlin: Springer-Verlag.

Arauzo-Carod, J., Liviano-Solis, D., & Manjón Antolín, M. (2010). Empirical Studies in Industrial Location: An Assessment of Their Methods and Results. *Journal of Regional Science, 50*, 685-711.

Arellano, M., & Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies, 58*(2), 277-297.

Aschauer, D. (1989). Is Public Expenditure Productive? *Journal of Monetary Economics, 23*(2), 177-200.

Barro, R. (1990). Government Spending in a Simple Model of Endogenous Growth. *The Journal of Political Economy, 98*(5), 103-111.

Barro, R., & Sala-i-Martin, X. (1992). Convergence. *The Journal of Political Economy, 100*(2), 223-251.

Barro, R., & Sala-i-Martin, X. (1992). Public Finance in Models of Economic Growth. *Review of Economic Studies, 59*(4), 645-661.

Barro, R., & Sala-i-Martin, X. (2004). *Economic Growth* (2nd ed.). New York: McGraw Hill.

Benjamin, D., Gunderson, M., & Riddell, C. (2005). *Labour Market Economics* (5th ed.). Toronto: McGraw-Hill Ryerson .

Bhattacharyay, B. (2009). Infrastructure Development for ASEAN Economic Integration. Working Paper No. 138, Asian Development Bank .

Biehl, D. (1991). The Role of Infrastructure in Regional Development. In R. Vickerman (Ed.), *Infrastructure and Regional Development* (pp. 9-35). London: Pion Limited.

Blair, J., & Premus, R. (1987). Major Factors in Industrial Location: A Review. *Economic Development Quarterly, 1*(1), 72-85.

Blanchard, O., Dell'Ariccia, G., & Mauro, P. (2010). Rethinking Macroeconomic Policy. IMF Staff Position Note SPN/10/03.

Blundell, R., & Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics, 87*(1), 115-143.

Blundell, R., & Bond, S. (2000). GMM Estimation with Persistent Panel Data: An Application to Production Functions. *Econometric Reviews, 19*, 321-340.

Boarnet, M. (1994). An Empirical Model of Intrametropolitan Population and Employment Growth. *Papers in Regional Science, 73*(2), 135-152.

Boarnet, M. (1998). Spillovers and the Locational Effects of Public Infrastructure. *Journal of Regional Science, 38*(3), 381-400.

Bollinger, C., & Ihlanfeldt, K. (2003). The Intra-urban Spatial Distribution of Employment: Which Government Interventions Make a Difference? *Journal of Urban Economics, 53*(3), 396-412.

Bond, S. R., Hoeffler, A., & Temple, J. (2001). GMM Estimation of Empirical Growth Models. Nuffield College Economics Working Paper No. 2001-W21.

Boscá, J., Escribá, F., & Murgui, M. (2002). The Effect of Public Infrastructure on the Private Productive Sector of Spanish Regions. *Journal of Regional Science, 42*(2), 301-326.

Bruinsma, F. (1995). The Impact of New Infrastructure on the Spatial Patterns of Economic Activities. Research Memorandum 1995-12.

Bruinsma, F., Rienstra, S., & Rietveld, P. (1997). Economic Impacts of the Construction of a Transport Corridor: A Multi-Level and Multi-Approach Case Study for the Construction of the A1 Highway in the Netherlands. *Regional Studies, 31*(4), 391-402.

Burgess, R., & Venables, A. (2004). Towards a Microeconomics of Growth. Policy Research Working Paper No. 3257, World Bank .

Button, K. (1998). Infrastructure Investment, Endogenous Growth and Economic Convergence. *The Annals of Regional Science, 32*(1), 145-162.

Buurman, J., & Rietveld, P. (1999). Transport Infrastructure and Industrial Location: The Case of Thailand. *Journal of the Applied Regional Science Conference, 11*(1), 45-62.

Calderón, C., & Serven, L. (2003). The Effects of Infrastructure Development on Growth and Income Distribution. Policy Research Working Paper No. 3400, World Bank .

Cameron, A., & Trivedi, P. (1998). *Regression Analysis of Count Data.* New York: Cambridge University Press.

Cameron, G., & Muellbauer, J. (1998). The Housing Market and Regional Commuting and Migration Choices. *Scottish Journal of Political Economy, 45*(4), 420-446.

Carroll, R., & Wasylenko, M. (1994). Do State Business Climates Still Matter? Evidence of A Structural Change. *National Tax Journal , 47*(1), 19-37.

Chandra, A., & Thompson, E. (2000). Does Public Infrastructure Affect Economic Activity? Evidence from the Rural Interstate Highway System. *Regional Science and Urban Economics, 30*(4), 457-490.

Chang, K., Hayakawa, K., & Matsuura, T. (2010). Location Choice of Multinational Enterprises in China: Comparison between Japan and Taiwan. Discussion Paper Series No. 251, Institute of Developing Economies.

Chasco, C., & López, A. (2010). Evolution of the Influence of Geography on the Location of Production in Spain (1930-2005). In A. Páez, J. Gallo, R. Buliung, & S. Dall'erba (Eds.), *Progress in Spatial Analysis: Methods and Applications* (pp. 407-440). Berlin: Springer.

Cheng, S. (2006). The Role of Labour Cost in the Location Choices of Japanese Investors in China. *Journal of Regional Science, 85*(1), 121-138.

Cheng, S. (2007). Structure of Firm Location Choices: An Examination of Japanese Greenfield Investment in China. *Asian Economic Journal, 21*(1), 47-73.

Chinitz, B. (1961). Contrast in Agglomeration: New York and Pittsburgh. *American Economic Review, 51*(2), 279-289.

Chung, T., Thang, D., Hoang, P., & Huy, N. (2010). Addressing the Bottlenecks: Towards an Effective Mechanism for Financing Infrastructure. Central Institute for Economic Management of Vietnam.

Clark, D., & Murphy, C. (1996). Countywide Employment and Population Growth: An Analysis of the 1980s. *Journal of Regional Science, 36*(2), 235-256.

Cohen, J., & Morrison, C. (2002). Public Infrastructure Investment, Inter-state Spatial Spillovers and Manufacturing Costs. *The Review of Economics and Statistics, 86*(2), 551-560.

Cohen, J., & Paul, C. (2004). Public Infrastructure Investment, Interstate Spatial Spillovers, and Manufacturing Costs. *The Review of Economics and Statistics, 86*, 551-560.

Crane, L., Burke, D., & Hanks, C. (1991). Highway District and Economic Sector Employment Effects of Transportation Expenditures. 190-200. Transportation Research Record No. 1305.

Dalenberg, D., & Partridge, M. (1995). The Effects of Taxes, Expenditures and Public Infrastructure on Metropolitan Area Employment. *Journal of Regional Science, 35*(4), 617-640.

Dalenberg, D., & Partridge, M. (1997). Public Infrastructure and Wages: Public Capital's Role as a Productive Input and Household Amenity. *Land Economics, 73*(2), 268-284.

Dalenberg, D., & Partridge, M. (1997). Public Infrastructure and Wages: Public Capital's Role as a Productive Input and Household Amenity. *Land Economics, 73*(2), 268-284.

Dalenberg, D., Partridge, M., & Rickman, D. (1998). Public Infrastructure: Pork or Jobs Creator? *Public Finance Review, 26*(1), 24-52.

Deichmann, U., Fay, M., Koo, J., & Lall, S. (2005). Agglomeration, Transport and Regional Development in Indonesia. Policy Research Working Paper No. 3477, World Bank.

Deichmann, U., Lall, S., Redding, S., & Venables, A. (2008). Industrial Location in Developing Countries. *The World Bank Research Observer, 23*(2), 219-246.

Delgado, M., & Álvarez, I. (2007). Network Infrastructure Spillover in Private Productive Sectors: Evidence from Spanish High Vapacity Roads. *Applied Economics, 39*(12), 1583-1597.

Devarajan, S., Swaroop, V., & Zou, H. (1996). The Composition of Public Expenditure and Economic Growth. *Journal of Monetary Economics, 37*(2), 313-344.

Duffy-Deno, K. (1998). The Effect of Federal Wilderness on County Growth in the Intermountain Western United States. *Journal of Regional Science, 381*(1), 109-136.

Duranton, G., & Puga, D. (2001). Nursery Cities: Urban Diversity, Process Innovation and the Life-Cycle of Products. *American Economic Review, 91*(5), 1454-1477.

Durlauf, S., Johnson, P., & Temple, J. (2005). Growth Econometrics. In P. Aghion, & S. Durlauf (Eds.), *Handbook of Economic Growth* (pp. 555-677). Elsevier.

Dutta, J., Horsewood, N., & Vasilakos, N. (2006). Industrial Activity in Indian States: The Role of Infrastructure. MPRA Paper No. 7184.

Easterly, W. (2003). National Policies and Economic Growth. In P. Aghion, & S. Durlauf (Eds.), *Handbook of Economic Growth* (pp. 1016-1053). Elsevier.

Ebert, R., & Stone, J. (1992). *Wage and Employment Adjustment in Local Labour Markets.* Michigan: W.E. Upjohn Institute for Employment Research.

Égert, B., Kozluk, T., & Sutherland, D. (2009). Infrastructure and Growth: Empirical Evidence. CESIFO Working Paper 2700.

Freidman, J. (1981). A Conditional Logit Model of the Role of Local Public Services in Residential Choice. *Urban Studies, 18*, 347-358.

Frenkel, J., & Romer, D. (1999). Does Trade Cause Growth? *American Economic Review, 89*(3), 379-399.

Fujita, M., & Thisse, J. (2002). *The Economics of Agglomeration: Cities, Industrial Location and Regional Growth.* Cambridge: Cambridge University Press.

Fujita, M., Krugman, P., & Venables, A. (1999). *The Spatial Economy: Cities, Regions and International Trade.* Cambridge, MA: MIT Press.

Futagami, K., Morita, Y., & Shibata, A. (1993). Dynamic Analysis of an Endogenous Growth Model with Public Capital. *Scandinavian Journal of Economics, 95*(4), 607-625.

Gabriel, S., & Rosenthal, S. (1989). Housing Location and Race: Estimates of A Multinomial Logit Model. *The Review of Economics and Statistics, 71*(2), 240-249.

García-Milà, T., & McGuire, T. (1992). The Contribution of Publicly Provided Inputs to States' Economies. *Regional Science and Urban Economics, 22*(2), 229-241.

García-Milà, T., McGuire, T., & Porter, R. (1996). The Effect of Public Capital in State-level Production Functions Reconsidered. *The Review of Economics and Statistics, LXXVIII*, 177-180.

Gebremariam, H., Gebremedhin, T., Schaeffer, P., T.T., P., & Jackson, R. (2011). Employment, Income, Migration and Public Services: A Simultaneous Spatial Panel Data Model of Regional Growth. *Journal of Regional Science, 91*(2), 275-297.

Gramlich, E. (1994). Infrastructure Investment: A Review Essay. *Journal of Economic Literature, XXXII*, 1176-1196.

Greenaway, D., Hine, R., & P.W., W. (1999). Exports, Export Composition and Growth. *Journal of International Trade and Economic Development, 8*(1), 41-51.

Greene, W. (2008a). *Econometric Analysis* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Greene, W. (2008b). Functional Forms for the Negative Binomial Model for Count Data. *Economics Letters, 99*(3), 585-590.

GSO. (2004). *Vietnam Statistical Data in the 20th Century, Volume 2.* Hanoi: Statistics Publishing House.

GSO. (2010). *Enterprises in Vietnam: 9 years at the Beginning of the 21th Century.* Hanoi: Statistics Publishing House.

GSO. (2011a). *Vietnam's Socioeconomic Conditions: 2001-2010.* Hanoi: Statistics Publishing House.

GSO. (2011b). *Migration and Urbanisation in Vietnam: Patterns, Trends and Differentials.* Hanoi: Vietnam General Statistics Office.

GSO. (2011c). *Age-Sex Structure and Marital Status of the Population in Vietnam.* Hanoi: Vietnam General Statistics Office.

Guimarães, P. (2008). The Fixed Effects Negative Binomial Model Revisited. *Economics Letters, 99*(1), 63-66.

Guimarães, P., Figueiredo, O., & Woodward, D. (2003). A Tractable Approach to the Firm Location Decision Problem. *Review of Economics and Statistics, 84*(1), 201-204.

Guimarães, P., Figueiredo, O., & Woodward, D. (2004). Industrial Location Modelling: Extending the Random Utility Framework. *Journal of Regional Science, 44*(1), 1-20.

Handerson, J. (1999). Overcome the Adverse Effects of Geography: Infrastructure, Health and Agricultural Policies. *International Regional Science Review, 22*(2), 233-237.

Hansen, E. (1987). Industrial Location Choice in San Paulo, Brazil: A Nested Logit Model. *Regional Science and Urban Economics, 17*(1), 89-108.

Hansen, N. (1965). Unbalanced Growth and Regional Development. *Western Economic Journal, 4*(1), 3-14.

Hashimzade, N., & Myles, G. (2010). Growth and Public Infrastructure. *Macroeconomic Dynamics, 14*(S2), 258-274.

Haughwout, A. (1999). State Infrastructure and the Geography of Employment. *State Infrastructure and the Geography of Employment , 30*(4), 549-566.

Haugwout, F. (1998). Aggregate Production Functions, Interregional Equilibrium and the Measurement of Infrastructure Productivity. *Journal of Urban Economics, 44*(2), 216-227.

Hausman, J., & McFadden, D. (1984). Specification Tests for the Multinomial Logit Model. *Econometrica, 52*(5), 1219-1240.

Hausman, J., Hall, B., & Griliches, Z. (1984). Econometric Models for Count Data with an Application to the Patents - R&D Relationship. *Econometrica, 52*(4), 909-938.

Hayter, R. (1997). *The Dynamics of Industrial Location: The Factory, the Firm and the Production System.* New York: Wiley.

Heiss, F. (2002). Structural Choice Analysis with Nested Logit Models. *The Stata Journal, 2*(3), 227-252.

Hensher, D., Rose, J., & Greene, W. (2005). *Applied Choice Analysis: A Primer.* Cambridge: Cambridge University Press.

Holl, A. (2004). Transport Infrastructure, Agglomeration Economies, and Firm Birth: Empirical Evidence from Portugal. *Journal of Regional Science, 44*(4), 693-712.

Holl, A. (2005). Start-ups and Relocations: Manufacturing Plant Location in Portugal. *Papers in Regional Science, 83*(4), 649-668.

Holtz-Eakin, D., & Schwartz, A. (1995). Spatial Productivity Spillovers from Public Infrastructure: Evidence from State Highways. *International Tax and Public Finance, 2*(3), 459-468.

Holzer, H., & Montgomery, E. (1990). Asymmetries and Rigidities in Wage Adjustments by Firms. Working Paper No. 3274, National Bureau of Economic Research.

Hotelling, H. (1929). Stability in Competition. *Economic Journal, 39*(153), 41-57.

Hulten, C., & Schwab, R. (1991). Public Capital Formation and the Growth of Regional Manufacturing Industries. *National Tax Journal, 44*(4), 121-134.

Irmen, I., & Kuehnel, J. (2009). Productive Government Expenditure and Economic Growth. *Journal of Economic Surveys, 23*(4), 692-733.

Islam, N. (1995). Growth Empirics: A Panel Data Approach. *Quarterly Journal of Economics, 110*(4), 1127-1170.

Jacobs, J. (1969). *The Economy of Cities.* New York: Random House.

Jiwattanakulpaisarn, P., Noland, R., Graham, D., & J.W., P. (2009a). Highway Infrastructure and State-level Employment: A Causal Spatial Analysis. *Journal of Regional Science, 88*(1), 133-159.

Jiwattanakulpaisarn, P., Noland, R., Graham, D., & J.W., P. (2009b). Highway Infrastructure Investment and County Employment Growth: A Dynamic Panel Regression Analysis. *Journal of Regional Science, 49*(2), 263-286.

Kelejian, H., & Prucha, I. (1998). A Generalised Spatial Two-Stage Least Squares Procedure Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. *Journal of Real Estate Finance and Economics, 17*(1), 99-121.

Kelejian, H., & Robinson, D. (1997). Infrastructure Productivity Estimation and Its Underlying Econometric Specifications: a Sensitivity Analysis. *Papers in Regional Science, 76*(1), 115-131.

Keller, W. (2002). Geographic Localisation of International Technology Diffusion. *The American Economic Review, 92*(1), 120-142.

Klapper, L., Amit, R., Guillén, M., & Quesada, J. (2007). Entrepreneurship and Firm Formation Across Countries. Policy Research Working Paper No. 4313, World Bank.

Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy, 99*(3), 483-499.

Krugman, P. (1993). First Nature, Second Nature, and Metropolitan Location. *Journal of Regional Science, 33*(2), 129-144.

Lall, S. (2007). Infrastructure and Regional Growth, Growth Dynamics and Policy Relevance for India. *The Annals of Regional Science, 41*(3), 581-599.

Lall, S., Koo, J., & Chakravorty, S. (2003). Diversity Matters: The Economic Geography of Industry Location in India. World Bank Policy Research Working Paper 3072.

Lee, K., Pesaran, M., & Smith, R. (1997). Growth and Convergence in a Multi-country Empirical Stochastic Solow Model. *Journal of Applied Econometrics, 12*, 357-392.

Levine, R., & Renelt, D. (1992). A Sensitivity Analysis of Cross-Country Growth Regressions. *American Economic Review, 82*(4), 942-963.

Li, S., & Park, S. (2006). Determinants of Locations of Foreign Direct Investment in China. *Management and Organization Review, 2*(1), 95-119.

Liu, N., Chen, Y., & Zhou, Q. (2007). Spatial Spillover Effects of Transport Infrastructure on Regional Economic Growth. *Journal of Southest University, 23*, 33-39.

Long, J. (1997). *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Saga Publications.

Long, J., & Fresse, J. (2001). *Regression Models for Categorical and Limited Dependent Variables Using Stata.* College Station, TX: Stata Press.

Lucas, R. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics, 22*(1), 3-42.

Mankiw, N., Rower, D., & D.N., W. (1992). A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics, 107*(2), 407-438.

Marshall, A. (1920). *Principles of Economics* (8th ed.). London: Macmillan and Co., Ltd.

Mas, M., Maudos, J., Perez, F., & Uriel, E. (1996). Infrastructure and Productivity in the Spanish Regions. *Regional Studies, 30*(7), 641-649.

McCann, P., & Shefer, D. (2004). Location, Agglomeration and Infrastructure. *Papers in Regional Science, 83*(1), 177-196.

McCarty, A. (1999). Vietnam's Labour Market in Transition. Paper presented at Law and Labour Market Regulation in Asia Conference, University of the Philippines, November 12, 1999.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behaviour. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105-142). New York: Academic Press.

McFadden, D. (1978). Modelling the Choice of Residential Location. In A. Karlqvist, L. Lundqvist, F. Snickars, & J. Weibull (Eds.), *Spatial Interaction Theory and Planning Models* (pp. 75-96). Amsterdam: North Holland.

McFadden, D., Tye, W., & Train, K. (1977). An Application of Diagnostic Tests for the Irrelevant Alternatives Property of the Multinomial Logit Model. *Transportation Research Record, 637*, 39-46.

Menon, J. (2009). Managing Success in Viet Nam: Macroeconomic Consequences of Large Capital Inflows with Limited Policy Tools. Working Paper Series on Regional Economic Integration No. 27, Asian Development Bank.

Mikelbank, B., & Jackson, R. (2000). The Role of Space in Public Capital Research. *International Regional Science Review, 23*(3), 235-258.

Milner, C., & Wright, P. (1998). Modelling Labour Market Adjustment to Trade Liberalisation in an Industrialising Economy. *Economic Journal, 108*(447), 509-528.

Modifi, A., & Stone, J. (1990). Do State and Local Taxes Affect Economic Growth? *Review of Economics and Statistics, 72*(4), 686-691.

Moreno, R., & López-Bazo, E. (2007). Returns to Local and Transport Infrastructure under Regional Spillovers. *International Regional Science Review, 30*(1), 47-71.

Morrison, C., & Schwartz, A. (1996). State Infrastructure and Productive Performance. *American Economic Review, 86*(5), 1095-1111.

Mota, I., & Brandão, A. (2012). The Determinants of Location Choice: Single Plants versus Multi-Plants. *Papers in Regional Science, 91*(4).

Munnell, A. (1990). How Does Public Infrastructure Affect Regional Economic Performance? *New England Economic Review, 1990*(9), 11-33.

Nadiri, M., & Mamuneas, T. (1994). The Effects of Public Infrastructure and R&D Capital on the Cost Structure and Performance of U.S. Manufacturing Industries. *Review of Economics and Statistics, 76*(1), 22-37.

Nadiri, M., & Mamuneas, T. (1998). *Contribution of Highway Capital to Output and Productivity Growth in the US Economy and Industries.* Report prepared for the Federal Highway Administration, Office of Transport Policy Studies.

Narayana, R., Kocherlakota, A., & Yi, K. (1995). Can convergence Regressions Distinguish between Exogenous and Endogenous Growth Models? . *Economics Letters, 49*, 211-215.

Nechyba, T., & Strauss, R. (1998). Community Choice and Local Public Services: A Discrete Choice Approach. *Regional Science and Urban Economics, 28*(1), 51-73.

Nefussi, B., & Schwellnus, C. (2010). Does FDI in Manufacturing Cause FDI in Business Services? Evidence from French Firm-level Data. *Canadian Economics Association, 43*(1), 180-203.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica, 49*(6), 1417-1426.

Nijkamp, P. (1986). Infrastructure and Regional Development: A Multidimensional Policy Analysis. *Empirical Economics, 11*(1), 1-21.

O'Connor, D. (1996). Labour Market Aspects of State Enterprise Reform in Vietnam. Working Paper No. 117, OECD Development Centre .

Owyong, D., & Thangavelu, S. (2001). An Empirical Study on Public Capital Spillovers from the USA to Canada. *Applied Economics, 33*(11), 1493-1499.

Parr, J. (2002). The Location Of Economic Activity: Central Place Theory and the Wider Urban System. In P. McCann (Ed.), *Industrial Location Economics* (pp. 32-82). Cheltenham: Edward Elgar.

Pereira, A. (2000). Is All Public Capital Created Equal? *Review of Economics and Statistics, 82*(3), 513-518.

Pereira, A., & Andraz, J. (2007). Public Investment in Transportation Infrastructures and Industry Performance in Portugal. *Journal of Economic Development, 32*(1), 1-20.

Pereira, A., & Andraz, J. (2010). On the Economic Effects of Public Infrastructure Investment: A Survey of the International Evidence. Working Paper No.108, College of William and Mary.

Pereira, A., & Roca-Sagalés, O. (2003). Spillover Effects of Public Capital Formation: Evidence from the Spanish Regions. *Journal of Urban Economics, 53*(2), 238-256.

Perrira, A., & Andraz, J. (2004). Public Highway Spending and State Spillovers in the USA. *Applied Economics Letters, 11*, 785-788.

Pierre, G. (2012). Recent Labour Market Performance in Vietnam through a Gender Lens. Policy Research Working Paper No. 6056, World Bank.

Porter, M. (2000). Location, Competition, and Economic Development: Local Clusters in a Global Economy. *Economic Development Quarterly, 14*(1), 15-34.

Preston, J. (2001). Integrating Transport with Socio-Economic Activity: A Research Agenda for the New Millennium. *Journal of Transport Geography, 9*(1), 13-24.

Pritchett, L. (1996). Mind Your P's and Q's: The Cost of Public Investment is Not the Value of Public Capital. World Bank Policy Research Working Paper 1660.

Prud'homme, R. (2004). *Infrastructure and Development.* Paper prepared for the Annual Bank Conference on Development Economics, The World Bank.

Quigley, J. (1985). Consumer Choice of Dwelling, Neighbourhood and Public Serivces. *Regional Science and Urban Economics, 15*, 41-63.

Quigley, J. (2008). Urbanisation, Agglomeration, and Economic Development. Commission on Growth and Development Working Paper No.19, World Bank.

Reynolds, P., Storey, D., & Westhead, P. (1994). Cross-national Comparisons of the Variation in New Firm Formation Rates. *Regional Studies, 28*(4), 443-456.

Rietveld, P. (1995). Infrastructure and Spatial Economic Development. *Annals of Regional Science, 29*, 117-119.

Rietveld, P., & Bruinsma, F. (1998). *Is Transport Infrastructure Effective?: Transport Infrastructure and Accessibility Impacts on the Space Economy.* Berlin: Springer-Verlag.

Rodrik, D., & Rodriguez, F. (1999). Trade Policy and Growth: A Sceptics' Guide to Cross National Evidence. NBER Working Paper 7081.

Romer, P. (1986). Increasing Returns and Long-run Growth. *Journal of Political Economy, 94*(5), 1002-1037.

Romp, W., & de Haan, J. (2007). Public Capital and Economic Growth: A Critical Survey. *Perspektiven der Wirtschaftspolitik, 8*(S1), 6-52.

Roodman, D. (2006). How to Do xtabond2: An Introduction to "Difference" and "System" GMM in Stata. Working Paper No. 103, Centre for Global Development .

Ros, J. (2000). *Development Theory and the Economics of Growth.* Ann Arbor: University of Michigan Press.

Sachs, J., & Warner, A. (1995). Economic Reform and the Process of Economic Integration. Brookings Papers on Economic Activity No. 1995(1).

Sala-i-Martin, X. (1997). I Just Ran Two Million Regressions. *Recent Empirical Growth Research, 87*(2), 184-188.

Seitz, H. (1993). A Dual Economic Analysis of the Benefits of the Public Road Network. *Annals of Regional Science, 27*(3), 223-239.

Seitz, H. (2000). Infrastructure, Industrial Development, and Employment in Cities: Theoretical Aspects and Empirical Evidence. *International Regional Science Review, 23*(3), 259-280.

Seitz, H., & Licht, G. (1995). The Impact of Public Infrastructure Capital on Regional Manufacturing Production Cost. *Regional Studies, 29*(3), 231-240.

Small, K., & Hsiao, C. (1985). Multinomial Logit Specification Tests. *International Economic Review, 26*(3), 619-627.

Solow, R. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics, 70*(1), 65-94.

Sridhar, K., & Wan, G. (2007). Firm Location Choice in Cities: Evidence from China, India, and Brazil. Research Paper No. 2007/56, UNU-WIDER .

Stephanedes, Y. (1990). Distributional Effects of State Highway Investment on Local and Regional Development. Transportation Research Record No. 1274, pp. 156-164.

Straub, S. (2008). Infrastructure and Growth in Developing Countries: Recent Advances and Research Challenges. Policy Research Working Paper No. 4460, World Bank.

Straub, S., & Terada-Hagiwara, A. (2010). Infrastructure and Growth in Developing Asia. Economics Working Paper Series No.231, Asian Development Bank.

Swan, T. (1956). Economic Growth and Capital Accumulation. *Economic Record, 32*(2), 334-361.

Tanzi, V., & Davoodi, H. (1997). Corruption, Public Investment, and Growth. Working Paper No. 1-23, International Monetary Fund.

Tobler, W. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography, 46*(2), 234-240.

Tokunaga, S., & Jin, S. (2011). Market Potential, Agglomeration and Location of Japanese Manufacturers in China. *Letters in Spatial Resource Science, 4*(1), 9-19.

Train, K. (2003). *Discrete Choice Methods with Simulation.* Cambridge, MA: Cambridge University Press.

United-Nations. (2011). *Re-establishing and Maintaining Macroeconomic Stability and Protecting the Poor, UN Country Team Statement for the Consultantive Group Meeting of June 2011.* Retrieved 09 02, 2012, from http://www.un.org.vn/en/feature-articles-press-centre-submenu-252/1841-re-establishing-and-maintaining-macroeconomic-stability-and-protecting-the-poor.html

Van Arkadie, B., & Mallon, R. (2004). *Vietnam: A Transition Tiger?* Canberra: Asia Pacific Press, Australian National University.

Venables, A. (2005). Spatial Disparities in Developing Countries: Cites, Regions and International Trade. *Journal of Economic Geography, 5*(1), 3-22.

Vickerman, R. (1980). *Spatial Economic Behaviour: The Microeconomic Foundations of Urban and Transport Economics.* London: Macmillan.

Vickerman, R. (1991b). Other Regions' Infrastructure in a Region's Development. In R. Vickerman (Ed.), *Infrastructure and Regional Development* (pp. 61-74). London: Pion.

Vickerman, R. (2002). Transport and Economic Development. Paris: OECD European Conference of Ministers of Transport, Round Table 119, pp. 139-177.

Vuong, Q. (1989). Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses. *Econometrica, 57*, 307-344.

Vuong, Q. (2010). *Financial Markets in Vietnam's Transition Economy: Facts, Insights, Implications.* Berlin: VDM Verlag Dr. Muller.

World-Bank. (1994). *World Development Report 1994: Infrastructure for Development.* Washington, DC: Oxford University Press.

World-Bank. (2000). *Private Solution for Infrastructure: Opportunities for Vietnam* (A Country Framework Report ed.). Washington D.C.: World Bank.

World-Bank. (2006). *Infrastructure Strategy: Cross-Sectoral Issues: Vietnam's Infrastructure Challenge* (A Country Framework Report ed.). Washington D.C.: World Bank.

Zhu, Y. (2011). Labour Markets in Vietnam: Developments under the Economic Reform and Globalisation. In J. Benson, & Y. Zhu (Eds.), *The Dynamics of Asian Labour Markets: Balancing Control and Flexibility* (pp. 148-164). London and New York: Routledge.

Zografos, K., & Stephanedes, Y. (1992). Impact of State Highway Investment on Employment along Major Highway Corridors. 151-155. Transportation Research Record No.1359.